
# CONTENT 2015

The Seventh International Conference on Creative Content Technologies

ISBN: 978-1-61208-392-6

March 22 - 27, 2015

Nice, France

**CONTENT 2015 Editors**

Hans-Werner Sehring, T-Systems Multimedia Solutions GmbH, Germany

Jaime Lloret Mauri, Polytechnic University of Valencia, Spain

# CONTENT 2015

# Forward

The Seventh International Conference on Creative Content Technologies (CONTENT 2015), held between March 22-27, 2015 in Nice, France, continued a series of events targeting advanced concepts, solutions and applications in producing, transmitting and managing various forms of content and their combination. Multi-cast and uni-cast content distribution, content localization, on-demand or following customer profiles are common challenges for content producers and distributors. Special processing challenges occur when dealing with social, graphic content, animation, speech, voice, image, audio, data, or image contents. Advanced producing and managing mechanisms and methodologies are now embedded in current and soon-to-be solutions.

The conference had the following tracks:

- Image and graphics
- Web content
- Content producers/distributors

Similar to the previous edition, this event attracted excellent contributions and active participation from all over the world. We were very pleased to receive top quality contributions.

We take here the opportunity to warmly thank all the members of the CONTENT 2015 technical program committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to CONTENT 2015. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the CONTENT 2015 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope CONTENT 2015 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of creative content technologies. We also hope that Nice, France provided a pleasant environment during the conference and everyone saved some time to enjoy the charm of the city.

**CONTENT 2015 Chairs**

**CONTENT Advisory Chairs**

Raouf Hamzaoui, De Montfort University - Leicester, UK
Jalel Ben-Othman, Université de Versailles, France
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Wolfgang Fohl, Hamburg University of Applied Sciences, Germany
Zhou Su, Waseda University, Japan

**CONTENT Industry/Research Chairs**

Ajith Abraham, Machine Intelligence Research Labs (MIR Labs), USA
Hans-Werner Sehring, T-Systems Multimedia Solutions GmbH, Germany
René Berndt, Fraunhofer Austria Research GmbH, Austria

**CONTENT Publicity Chairs**

Lorena Parra, Universidad Politécnica de Valencia, Spain
Samuel Kosolapov, Braude Academic College of Engineering, Israel
Wilawan Inchamnan, Queensland University of Technology, Australia
Javier Quevedo-Fernandez, Eindhoven University of Technology, The Netherlands

# CONTENT 2015

## Committee

### CONTENT Advisory Chairs

Raouf Hamzaoui, De Montfort University - Leicester, UK
Jalel Ben-Othman, Université de Versailles, France
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Wolfgang Fohl, Hamburg University of Applied Sciences, Germany
Zhou Su, Waseda University, Japan

### CONTENT Industry/Research Chairs

Ajith Abraham, Machine Intelligence Research Labs (MIR Labs), USA
Hans-Werner Sehring, T-Systems Multimedia Solutions GmbH, Germany
René Berndt, Fraunhofer Austria Research GmbH, Austria

### CONTENT Publicity Chairs

Lorena Parra, Universidad Politécnica de Valencia, Spain
Samuel Kosolapov, Braude Academic College of Engineering, Israel
Wilawan Inchamnan, Queensland University of Technology, Australia
Javier Quevedo-Fernandez, Eindhoven University of Technology, The Netherlands

### CONTENT 2015 Technical Program Committee

Marios C. Angelides, Brunel University - Uxbridge, UK
Kambiz Badie, Research Institute for ICT & University of Tehran, Iran
David Banks, University of Tennessee, USA
Rene Berndt, Fraunhofer Austria Research GmbH, Austria
Christos Bouras, University of Patras and Computer Technology Institute & Press «Diophantus», Greece
Oliver Bown, University of Sydney, Australia
Andrew Brown, Queensland University of Technology, Australia
Yiwei Cao, RWTH Aachen University, Germany
Wojciech Cellary, Poznan University of Economics, Poland
Lijun Chang, University of New South Wales, Australia
Savvas A. Chatzichristofis, Democritus University of Thrace, Greece
Chi-Hua Chen, National Chiao Tung University, Taiwan, R.O.C.
Octavian Ciobanu, "Gr.T. Popa" University of Medicine and Pharmacy – Iasi, Romania

Wojciech R. Wiza, Poznan University of Economics, Poland
Shigang Yue, University of Lincoln, UK
Juan Zamora, Universidad Técnica Federico Santa María, Chile

**Copyright Information**

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

**Table of Contents**

# Time Coherent Animation of Dynamic 3D Objects from Kinect Camera using 3D Features

Naveed Ahmed
Department of Computer Science
University of Sharjah
Sharjah, United Arab Emirates
nahmed@sharjah.ac.ae

Salam Khalifa
Department of Computer Science
University of Sharjah
Sharjah, United Arab Emirates
salam.khalifa@gmail.com

***Abstract*— We present a system for creating a time coherent animation of dynamic 3D objects from Kinect sensor using 3D features. We record a dynamic object using the Kinect sensor resulting in a sequence of Red, Green, Blue and Depth (RGB-D) frames of that object. We extract 3D features from the depth data and using these features, we estimate the motion of unrelated 3D geometry between two consecutive frames. Extending this motion compensation over the complete sequence, we manage to create a time coherent 3D animation of a dynamic 3D object, which can be used in a number of applications that require smooth temporal data for post capture analysis, e.g., action or object recognition, motion capture, 2D or 3D gesture recognition, motion editing, or non-interactive 3D animation.**

***Keywords-Kinect; Dynamic Point Clouds; Point Sampling; 3D Animation***

## I. INTRODUCTION

Kinect sensor [13] has opened a completely new way of how we can visualize real world objects in the digital domain. It not only allows capturing the general appearance of a real-world object by means of a Red, Green, and Blue (RGB) camera, but also its 3D geometric shape using the depth camera. The latter feature allows it to be used in a number of applications that rely on capturing the true appearance of any object or gesture-based natural user interface. When used for capturing the dynamic 3D geometry, Kinect captures a depth image for each frame of the recorded video. The depth image provides the distance from the object to the camera. This depth image can be resampled into a 3D point cloud, in which every point has a specific location in the 3D space. If one has to visualize the dynamic 3D geometry of a real world object then this dynamic 3D point cloud representation is sufficient. Unfortunately, this representation is not suitable for a number of post processing tasks, e.g., motion analysis, motion compression, action recognition, etc. The reason being that each frame of the dynamic 3D point cloud is completely independent of the other, and there is no time coherence in the data. If this data has to be used in any of the aforementioned applications, then this time coherence has to be established, by either some form of mapping between each frame of the animation or estimating the motion of the dynamic object over the whole sequence.

In this paper, we present a new method for creating time coherent 3D animation from a sequence of depth images obtained from a Kinect camera. Our method is not confined to the depth image representation, rather we resample the depth images into dynamic 3D point clouds, and therefore our approach will work for any 3D animation in the form of 3D point clouds. We show that by means of 3D feature extraction from 3D point clouds, we can estimate motion of a dynamic object between two consecutive frames. Tracking the object over the whole sequence, results in a compact motion compensation representation of a time coherent 3D animation. As the result of our work, we do not need to store a dynamic object at each frame; rather the motion of the dynamic object is encoded at each subsequent frame. Thus, our method smoothly tracks one dynamic 3D object over the whole sequence that goes through the same animation that was captured in the non-coherent representation. The resulting time coherent representation can be employed in a number of applications for a post processing and post recording analysis.

A number of methods has been proposed to create 3D animations. Carranza et al. [7] presented a Free-viewpoint system for 3D animation that recorded a dynamic object using eight RGB cameras. The video data from eight cameras was used to capture the motion and shape of the actor. They used a template model to capture the approximate shape of the real-world actor. They did not use the multi-view data to reconstruct the dynamic 3D geometry. One of the drawbacks of this work was that it did not capture the surface reflectance of dynamic objects. Theobalt et al. [16] extended the work of Carranza et al. [7] and added dynamic surface reflectance estimation. Still, the method used a course template mesh, which was not the true geometry reconstructed from the video data. Different approaches were employed by Vlasic et al. [17] and Aguiar et al. [8] to capture realistic shape appearance and motion of dynamic objects. For the shape, both methods relied on high quality laser scans of the person. Both methods used high definition RGB cameras to capture the appearance of the actor. They differ in their approach on capturing the dynamic shape. Vlasic et al. [17] used a skeleton-based

 ISBN: 978-1-61208-392-6 

method to deform the template geometry. Their method works well for most of the objects, but had some limitations for the objects where the skeleton representation does not apply. On the other hand, Aguiar et al. [8] used a data driven approach to capture shape deformations. Their method is well suited to any kind of 3D geometry representation, as long as a high quality template model is available. Ahmed et al. [1] presented a work on reconstructing time coherent 3D animation without using the template geometry. They first created dynamic 3D visual hulls of the real world object, which were then tracked using a feature based dense matching algorithm. Their feature based matching does not incorporate any geometric features. They obtain Scale Invariant Feature Transform (SIFT) features from RGB images and match them over the animation while mapping them on the visual hulls. Their method is not suitable for 3D point cloud representation because it explicitly requires a smooth 3D surface to be available for calculating the geodesic distance between two points. Whereas, our method does not rely on any surface information because getting a smooth surface representation from noise Kinect data is a very challenging task.

In the past four or five years, an increasing number of methods for 3D geometry and animation reconstruction have started using depth cameras. It was initially made possible with the availability of the relatively low cost Time-of-Flight [11] depth cameras that can provide low-resolution dynamic geometry at high frame rate. A number of applications were proposed [3] [11] using the Time-of-Flight cameras. Microsoft completely changed the landscape of a general-purpose depth camera by bringing the extremely low cost Kinect for Xbox 360 as a general consumer electronics equipment. Kinect was a revolutionary device, because it could capture both color and depth data at 30 frames per second. The resolution of both cameras is really low (640x480) but because of its lower cost and deployment with Xbox 360, it was widely adopted. Apart of the gaming community, the research community also employed Kinect in a number of applications. A number of new methods were proposed in the areas of gesture recognition, motion capture, surface deformation, and motion editing.

Researchers have been employing depth cameras for reconstructing both dynamic and static real-world objects. One or more depth cameras were used by Kim et al. [11] and Castaneda et al. [6] for reconstructing a three-dimensional representation of static objects. Depth cameras are also used to reconstruct 3D shape, pose, and motion in the works presented by Berger et al. [4], Girshich et al. [9], Weiss et al. [18], and Baak et al. [3]. Multiple depth sensors are employed for capturing the dynamic scenes. Ahmed et al. [19] used six Kinect sensors to record a dynamic 3D object and create a 3D animation. Kim et al. [10] and Berger et al. [4] also used multiple depth sensors for object acquisition. Both of these methods do not establish any time coherence in the time varying data. On the other hand, Ahmed et al. [19] do reconstruct the time-coherent 3D animation but their work relies on RGB data for the feature points, whereas we show that one can reconstruct time coherent animation only using the geometric features.

Our work derives from the motivation of not using RGB data in the time coherent animation reconstruction. Even though RGB data has been successfully used in this line of research, it requires an additional mapping from depth data to RGB. In case of Kinect, this mapping is only one directional, i.e., from depth to RGB and that too is many to one. It means that multiple depth values can be mapped to a single RGB pixel. Thus, a feature point in RGB has an ambiguous representation in the three-space geometry. We therefore propose a framework that can work on the acquisition by one or more depth cameras and only utilizes the depth data for time coherent 3D animation reconstruction. We record a sequence using the Kinect camera and resample it in the form of a dynamic 3D point cloud representation. These point clouds are not time coherent and are completely independent of each other. In the following step, we extract a number of 3D features from each point cloud and match two consecutive frames using these features, starting from the first frame of the sequence. Using the mapping between the first two frames, we estimate the motion of the 3D point cloud between the two frames. This tracking is done over the whole sequence and we end up with a representation where for each frame we only need to store the motion with respect to the previous frame. Thus, our main contribution is a motion compensation representation by means of tracking using 3D features that creates a time coherent animation of a dynamic object.

Figure. 1 One RGB and depth frame captured from Kinect.

## II. VIDEO ACQUISITION

We acquire the dynamic 3D object using one Microsoft Kinect camera. Our method is not limited to a single camera setup, and can easily be extended to data from multiple cameras, as long as it is registered in a global coordinate space. An example of a multi-view setup can be seen in the work of Ahmed et al. [19].

The Kinect can capture two simultaneous video data streams, one RGB stream, and one depth stream. We use Microsoft Kinect Software Development Kit (SDK) to capture RGB-D data. For our method, we do not use the RGB stream but we capture to verify the acquisition process and make sure that the results are consistent with the depth stream. Kinect SDK can record both streams at different resolutions. At 30 frames per second, it can only record at 640x480 or a lower resolution. It can also capture at 1024x768 but the frame rate drops to 15 frames per second. Since we are interested in recording a dynamic object, the frame rate gets higher preference than the resolution. Therefore, we record both streams at 640x480 at 30 frames per second. The recording is stored in a high-speed memory buffer to avoid any input/output (IO) read/write overhead during the process. Once the recording is finished, each frame of the captured data is written to the disk.

The acquisition setup provides us with a frame-by-frame sequence of both RGB and depth data. One RGB and depth frame of the captured sequence can be seen in Fig. 1.

## III. CALIBRATION AND BACKGROUND SUBTRACTION

Our acquisition system provides us with both RGB and depth streams. Each stream is comprised of a sequence of frames. For example, each frame of the depth stream is an intensity image of the resolution 640x480, where each pixel is associated to a depth value. There is no notion of how these depth values will be mapped to the three space for the visualization. Similarly, there is no relationship between the depth and RGB stream. For some methods that need both RGB and depth streams, a mapping has to be established between them.

For our work, we need two types of calibrations. First, we need to estimate the intrinsic parameters of the depth camera. Then, we need to find the mapping of the depth values provided by Kinect in a form of a two-dimensional depth image in the three-dimensional world coordinate space. Optionally, we also obtain the mapping between the RGB and depth stream to verify the correct acquisition of the data.

We use Matlab Camera Calibration toolkit [23] for the intrinsic calibration. We record a checkerboard from both color and infrared sensors to facilitate this calibration. We use the tool Kinect RGB Demo by Nicola Burrus [5] to convert depth data to real world three space distances, and find the mapping between RGB and depth streams. The depth camera calibration allows us to resample a depth image into a 3D point cloud and the RGB and depth stream mapping allows us to visualize the resampled point cloud with the color information to validate our acquisition setup. An example of the resampled 3D point cloud and the mapping of RGB to depth can be seen in Fig. 2.

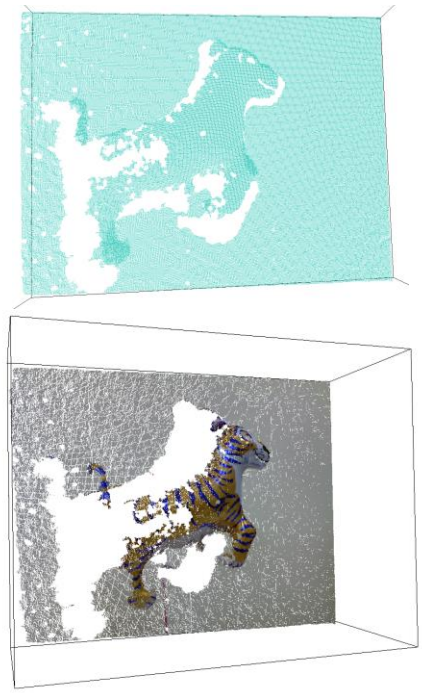


Figure. 2 A resampled 3D point cloud can be seen in the top image, while the same point cloud with RGB colors using the RGB and depth mapping can be seen in the bottom image.

We store 3D point clouds in the Point Cloud Data (PCD) file format using the Point Cloud Library [14]. The Point Cloud Library allows for efficient storage and manipulation of the point cloud data. It also has a number of algorithms implemented that can be used to analyze the point cloud data. We make an extensive use of this library in our work that will be explained in the next section.

After obtaining the resampled point cloud, we perform background subtraction to separate the dynamic object from the static background. In the first step, we record the same scene without the dynamic object. For the background subtraction, we record 30 frames of the background. Afterward, we average the 30 depth frames to average out the noise in the data. The mean background depth image is then subtracted from each depth frame of the recorded video sequence. This results in a separation of the dynamic model from the background and significantly reduces the storage cost for the point cloud. The depth data from Kinect is marred by very high temporal noise. This is a limitation of the technology and because of the high frame rate, it can be really pronounced when visualized. Therefore, using the

 ISBN: 978-1-61208-392-6 

Point Cloud library we also de-noise the data by means of simple Gaussian filtering. A 3D point cloud after background subtraction and filtering can be seen in Fig. 3.

## IV. TIME COHERENT ANIMATION

So far, we have obtained a segmented 3D point cloud for each frame of the video sequence. These point clouds are completely independent of each other and there is no coherence from one frame to the other. This is the preliminary requirement of data representation for our method to create time coherent 3D animation. Our method is not limited to data obtained from Kinect. As long as dynamic 3D point cloud data is available, from either depth or RGB cameras, our method will work equally well. The only reason we are using Kinect is that we can obtain dynamic 3D point cloud representation from just one camera, whereas in a traditional RGB camera acquisition system, at least two cameras are required to reconstruct the depth information.

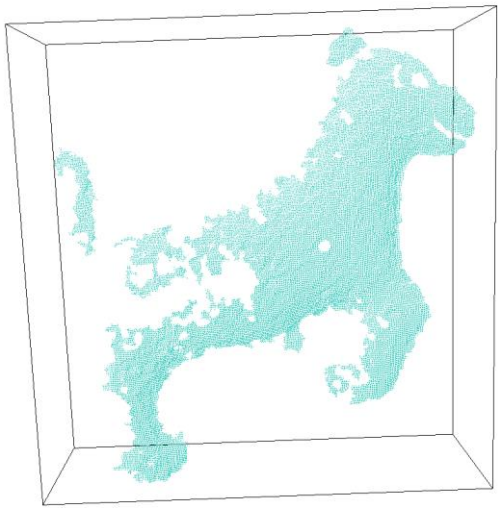


Figure. 3 Result from the background subtraction. The bounding box is significantly reduced compared to the original point cloud in Fig. 2.

To reconstruct time coherent animation we start by estimating a mapping between two consecutive frames of the dynamic scene sequence. We start by extracting 3D features from the first two frames $t_0$ and $t_1$. These features are then matched to find a sparse mapping between the two frames. This sparse matching is used to estimate the motion between the two frames. If the object undergoes a simple motion, e.g., translation, then only one match between the two frames is sufficient to track the point cloud from one frame to the next. Three or more matches can estimate a rigid body transform. On the other hand, if the motion is non-linear, which is true in our recordings then we need to find the motion of every point in the point cloud. We estimate the motion of all the points in the point cloud by using the sparse matching as the starting point. In the subsequent steps, we track $t_0$ over the whole sequence, resulting in a time coherent animation. Thus, our time coherent animation reconstruction algorithm takes the following form:

1) *Find 3D feature points at each frame*
2) *Match two consecutive frames starting from $t_0$ & $t_1$*
3) *Estimate motion of each point on $t_0$*
4) *Using the estimated motion at t0, track it to $t_1$*
5) *Loop from step 2 and track $t_0$ over the sequence*

In the first step, we find a number of 3D feature points for each frame of the 3D point cloud. We use the Point Cloud Library to estimate the following 3D features:

1) Estimate 3D SIFT over the depth image. The depth image is treated as an intensity image, and every feature point has a unique three-space location [20].
2) For every point on the point cloud, we estimate its underlying curvature and normal.
3) Using the normal information from step 2, estimate Clustered Viewpoint Feature Histogram (CVFH) descriptor [21].

These 3D features are then used to find a sparse correspondence between $t_0$ and $t_1$. 3D SIFT features are matched over the two depth images. It provides us with a one to one mapping for a sparse number of 3D positions. While matching 3D SIFT descriptors, we make use of curvature and normal to ensure that the matching is not an outlier. On the other hand, CVFH provides us with the matching clusters. Sparse matching approach is incorporated in earlier works, e.g., [1] and [22]. Our method is significantly different from those works, because it is incorporating 3D features. In order to find the one to one matching from the sparse correspondences, we make use of the approach from Salam et al. [22]. The one to one matching from 3D SIFT allows us to estimate the motion vector for the sparse matching points:

$$M_S = FP_1 - FP_0 \quad (1)$$

where $\boldsymbol{M_S}$ is the set of motion vectors for all 3D SIFT feature points. $\boldsymbol{FP_1}$ and $\boldsymbol{FP_0}$ are the feature points at frame 1 and 0 respectively. Similarly, for each cluster from CVFH we estimate its motion vector:

$$M_C = CP_1 - CP_0 \quad (2)$$

where $\boldsymbol{M_C}$ is the set of motion vectors for all clusters. $\boldsymbol{CP_1}$ and $\boldsymbol{CP_0}$ are the centroids of the clusters at frame 1 and 0 respectively. In the next step, we need to estimate the motion of all the points at $t_0$. For every point at $t_0$, we find the four nearest points in $\boldsymbol{FP_0}$ and the nearest cluster with respect to its centroid $\mathbf{CP_0}$. Each of these nearest points and cluster has an associated motion vector, i.e. $\mathbf{M_{S0}}$, $\mathbf{M_{S1}}$, $\mathbf{M_{S2}}$, $\mathbf{M_{S3}}$, and $\mathbf{M_{C0}}$. The motion vector for any point at $t_0$ is then defined as:

$$M_V = (M_{S0} + M_{S1} + M_{S2} + M_{S3} + M_{C0})/5 \quad (3)$$

where $\boldsymbol{M_v}$ is the average motion vector for the 3D point. Once this motion vector for each point is established, it is used to track $t_0$ to $t_1$. Thus, for the time step $t_1$ we do not need to store the complete point cloud, rather we can represent $t_0$ at $t_1$ in a motion compensation representation.

Using the estimated motion $t_0$ and $t_1$, we trivially track $t_0$ over the whole sequence. For example, in the next step the mapping between $t_1$ and $t_2$ is established but this mapping is used to find the motion vector of each point of tracked $t_0$. The same procedure then follows for all subsequent frames.

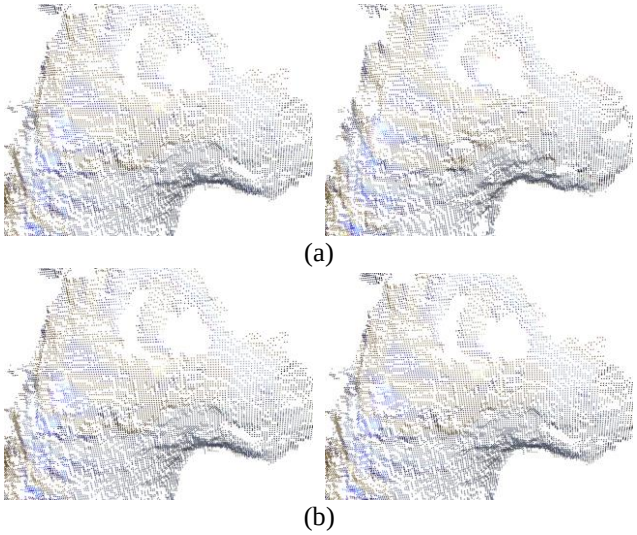
## V. RESULTS

We use two types of data sets to validate our method. Both data sets are acquired through a single Kinect and each is 100 frames long. In the first sequence, we only have one object in the scene whereas in the second sequence there are two dynamic objects. Our method for creating time coherent animation managed to track both sequences completely. The result of the animation is a single point cloud tracked over the whole sequence.

Our method is very efficient in its implementation. On average we can track 10 frames each second, thus tracking a 100 frames animation takes less than 2 minutes on a Core i5 2.4 Ghz processor. Some results of our tracking method can be seen in Fig. 4.

Our method is subject to some limitations. One of the major limitations is the quality of the data. Depending on the speed of the motion, the number of 3D features can decrease, which will result in low quality of time coherent animation. Even using RGB images for detecting feature points will not solve this problem because fast motion introduces motion blur, which reduces the quality of the RGB data. This limitation can be rectified by using high frame-rate cameras. Other limitation is the choice of 3D features. We are limited to the types of 3D features because of our data representation. Most of the 3D features require a surface representation. In principal, one can generate the surface from a 3D point cloud. For the data from Kinect, it is a difficult problem because the depth data from Kinect has a very high temporal noise, which makes surface estimation a research problem in itself. In future, we would like to simulate a smoother point cloud and test surface reconstruction on it and evaluate the results from different types of 3D features.

Despite the limitations, we show that it is possible to create time coherent animation from dynamic 3D point clouds from Kinect using only the 3D features from the depth data.



(a)

(b)

Figure. 4 Two non-coherent consecutive frames of 3D point cloud are shown in (a). Whereas (b) shows the same two frames generated using time coherent animation method. The frames are frame #0 and frame #80. It can be seen that in the non-coherent point clouds (a), the points change between the frames, esp. the effect is visible in the shape of the eye and around. The point cloud at frame #0 is tracked to frame #80 and does not show any changes in its shape by frame #80 (b).

## VI. CONCLUSION

We presented a method to create time coherent animation from dynamic 3D point clouds using only 3D features from the depth data. We show that noisy data from a Kinect camera can be resampled to create a dynamic 3D point cloud representation of a dynamic object. After the internal calibration and background subtraction, we manage to isolate the dynamic object for creating a time coherent animation. Our time coherent animation reconstruction method is an iterative process, which uses 3D features from the point cloud to match two consecutive frames. The initial matching is propagated from first frame to the last resulting in a time coherent animation where a single 3D point cloud is tracked over the complete sequence. Our method is not restricted to the data obtained from Kinect. It can work for any animation as long as it is represented in the form of dynamic 3D point clouds. In future, we plan to extend our work to incorporate dynamic surface reconstruction and new 3D feature representations. The resulting time coherent animation from our method can be used in a number of applications, e.g., action or object recognition, gesture recognition, motion capture, analysis and compression.

## ACKNOWLEDGMENT

This work was supported by the seed grant (110226) for new faculty members from the Department of Graduate Studies, University of Sharjah, UAE.

 ISBN: 978-1-61208-392-6

## REFERENCES


[1] N. Ahmed, C. Theobalt, C. Rossl, S. Thrun, and H. P. Seidel, "Dense Correspondence Finding for Parametrization-free Animation Reconstruction from Video," Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 08), IEEE Press, June 2008, pp. 1-8, doi:10.1109/CVPR.2008.4587758.

[2] N. Ahmed, "A System for 360 degree Acquisition and 3D Animation Reconstruction using Multiple RGB-D Cameras," Proc. Computer Animation and Social Agents (CASA 2012), May 2012, pp. 1-4.

[3] A. Baak, M. Muller, G. Bharaj, H. P. Seidel, and C. Theobalt, "A Data-driven Approach for Real-time Full Bodypose Reconstruction from a Depth Camera," Proc. IEEE International Conference on Computer Vision (ICCV 2011), November 2011, pp. 1092-1099, doi:10.1109/ICCV.2011.6126356.

[4] K. Berger, K. Ruhl, Y. Schroeder, C. Bruemmer, A. Scholz, and M. A. Magnor, "Markerless Motion Capture Using Multiple Color-depth Sensors," Proc. 16th International Workshop on Vision Modeling and Visualization (VMV 2011), October 2011, pp. 317–324.

[5] N. Burrus, "Kinect RGB Demo," http://www.computervisiononline.com/software/kinect-rgbdemo, last access 07/03/2015.

[6] V. Castaneda, D. Mateus, and N. Navab, "Stereo Time-of-Flight," Proc. IEEE International Conference on Computer Vision (ICCV 2011), November 2011, pp. 1684-1691, doi:10.1109/ICCV.2011.6126431.

[7] J. Carranza, C. Theobalt, M. Magnor, and H. P. Seidel, "Free-viewpoint Video of Human Actors," Proc. ACM Siggraph, August 2003, pp. 569-577, doi:10.1145/882262.882309.

[8] E. D. Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H. P. Seidel, and S. Thrun, "Performance Capture from Sparse Multi-view Video," Proc. ACM Siggraph, August 2008, pp. 15-23, doi:10.1145/1399504.1360697.

[9] R. Girshick, J. Shotton, P. Kohli, and A. C. Fitzgibbon, "Efficient Regression of General-activity Human Poses from Depth Images," Proc. IEEE International Conference on Computer Vision (ICCV 2011), November 2011, pp. 415-422, doi:10.1109/ICCV.2011.6126270.

[10] Y. M. Kim, D. Chan, C. Theobalt, and S. Thrun, "Design and Calibration of a Multi-view ToF Sensor Fusion System," Proc. IEEE Computer Vision and Pattern Recognition Workshop on Time-of-Flight, June 2008, pp. 55-62.

[11] Y. M. Kim, C. Theobalt, J. Diebel, J. Kosecka, B. Micusik, and S. Thrun, "Multi-view Image and ToF sensor Fusion for Dense 3D Reconstruction," Proc. IEEE Workshop on 3-D Digital Imaging and Modeling (3DIM09), September 2009, pp. 1542-1549, doi:10.1109/ICCVW.2009.5457430.

[12] D. G. Lowe, "Object Recognition from Local Scale Invariant Features," Proc. IEEE International Conference on Computer Vision (ICCV), 1999, pp. 1150–1157, doi:10.1109/ICCV.1999.790410.

[13] Microsoft, "Kinect for microsoft windows and xbox 360," http://www.kinectforwindows.org/, last access 07/03/2015.

[14] B. R. Radu, and C. Steve, "3D is here: Point Cloud Library (PCL)," IEEE International Conference on Robotics and Automation (ICRA), May 2011, pp. 1-4, doi:10.1109/ICRA.2011.5980567.

[15] A Tevs, A. Berner, M. Wand, I. Ihrke, and H. P. Seidel, "Intrinsic Shape Matching by Planned Landmark Sampling," Proc. 32nd International Conference of the European Association for Computer Graphics (Eurographics), April 2011, pp. 543-552.

[16] C. Theobalt, N. Ahmed, G. Ziegler, and H. P. Seidel, "High-Quality Reconstruction of Virtual Actors from Multi-view Video Streams," IEEE Signal Processing Magazine, vol 24, 2007, pp. 45-57.

[17] D. Vlasic, I. Baran, W. Matusik, and J. Popovic, "Articulated Mesh Animation from Multi-view Silhouettes," Proc. ACM Siggraph, August 2008, pp. 24-31, doi: 10.1145/1399504.1360696.

[18] A. Weiss, D. Hirshberg, and M. J. Black, "Home 3D Body Scans from Noisy Image and Range Data," Proc. IEEE Internation Conference on Computer Vision (ICCV), November 2011, pp. 1951-1958, doi:10.1109/ICCV.2011.6126465.

[19] N. Ahmed, and I. Junejo, "A System for 3D Video Acquisition and Spatio-Temporally Coherent 3D Animation Reconstruction using Multiple RGB-D Cameras," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 6, 2013, pp. 113-129.

[20] D. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol 60, 2004, pp. 91-110.

[21] A. Aldoma, N. Blodow, D. Gossow, S. Gedikli, R.B. Rusu, M. Vincze and G. Bradski, "CAD-Model Recognition and 6 DOF Pose Estimation," Proc. IEEE International Conference on Computer Vision Worskhops, 2011, pp. 585-592, doi:10.1109/ICCVW.2011.6130296.

[22] S. Khalifa, and N. Ahmed, "Temporally Coherent 3D Animation Reconstruction from RGB-D Video Data", Proc. International Conference on Computer Graphics, Visualization, Computer Vision and Image Processing (ICCGCVIP), 2014, pp. 1383-1389.

[23] J. Y Bouguet, "Camera Calibration Toolbox for Matlab," http://www.vision.caltech.edu/bouguetj/calib_doc/, last access 07/03/2015.


 ISBN: 978-1-61208-392-6 

# Estimating the Mood of Paintings from Color Combinations using Machine Learning

Dongwann Kang and Kyunghyun Yoon

School of Computer Science & Engineering
Chung-Ang University
Seoul, Korea
Email: {dongwann,khyoon}@cglab.cau.ac.kr

***Abstract*—A color image scale is a useful tool that enables designers to express mood through color combinations. This paper proposes a method for estimating the mood in color image scales from three-color combinations using machine learning. First, we find the relation between the mood and the properties of the color combinations. Then, we extract the three most dominant colors from the image. Finally, we estimate the mood of the painting via the properties of the three dominant colors extracted.**

***Keywords–color image scale;mood; color combinations; painting.***

## I. INTRODUCTION

Color is psychologically perceived by humans. Certain colors intuitively evoke certain moods or feelings in many people. Most artists even intentionally use color to convey its own meaning. In numerous studies, the colors in images have been determined to be an important factor that affects mood [1].

A Color Image Scale [2] is used extensively as a tool for selecting colors while considering mood in various fields, such as product design and cloth coordination. A color image scale consists of two axes: warm/cool and soft/hard. The positions of single color, color combinations, paintings, potteries, buildings, and moods on color image scales have been determined from user studies. For example, Figure 1 shows the positions of various three-color combinations on a color image scale. In this paper, we assume that an item in one position on the color image scale is strongly related to any other in the same position. The objective of this paper is to find moods from paintings by considering this underlying assumption.

The remainder of this paper is organized as follows. In Section II, we explain our approach for finding the correlation between colors and mood. In Section III, we present our proposed method for estimating the mood of a painting using correlation and discuss the results obtained. Finally, we conclude this paper with a summary of our ideas and an outline of future work in Section IV.

## II. ESTIMATING MOOD FROM THREE COLOR COMBINATIONS

In our estimation of moods from paintings, we use three-color combinations surveyed by Kobayashi [3]. Kobayashi provides three-color combinations tagged as moods, such that the mood of a painting can be estimated via three-color combinations extracted from it. The name of each color combination and the position of the corresponding mood in the color image scale are also provided. Because the positions of mood

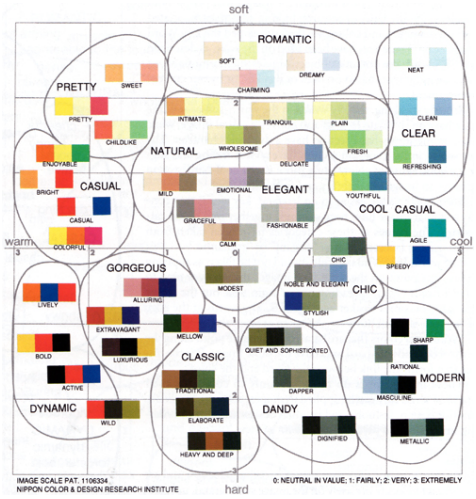



Figure 1. Three-color combinations on a color image scale.

keywords are graphically represented in [3], we estimated the position each mood by acquiring the center position of text in a graph, also obtained from [3]. Consequently, we obtained three-color combinations, with the names of the three colors, the name of each mood tagged onto each combination, and the position of the moods in the color image scale.

Although Kobayashi provided a number of three-color combinations, each tagged with its own mood, not all the three-color combinations available in paintings are represented. Thus, in order to be able to estimate mood from any three random colors, determination of the relation between each color in Kobayashis three-color combination is very important. To estimate this relation, we employed a machine learning technique. First, we extracted features from the colors in each combination, such as the hue/saturation/luminance difference between two colors and the average hue/saturation/luminance value of the three colors. Consequently, we obtained a 12-dimensional feature for each three-color combination. Next, we generated data pairs, with the features and the two-dimensional position of the mood tagged onto the data, for each three-color combination. Finally, we derived a prediction function that is able to estimate the coordinates of the mood of random three-color combinations using linear regression [4].

We then conducted an experiment using 936 three-color combinations and 174 moods. In order to ignore the order of colors in a combination, we generated all possible combinations of the given 936 three-color combinations, resulting

 ISBN: 978-1-61208-392-6

in six combinations being generated from each three-color combination. The range of both coordinates of each color image scale was $[-3 : +3]$. In our experiment, the prediction error magnitude was recorded as $0.64$. In further analysis, the significant factors appeared to be average hue, hue difference, average saturation, and intensity.

## III. Estimating Mood from Dominant Colors of Image

In the previous section, we discussed and predicted moods from three-color combinations. Consequently, we hypothesized that if we obtained three-color combinations from a painting, then we should be able to estimate the mood of that painting. In this paper, we assume that the three colors that are most dominant in a painting affect the mood of the painting in a manner similar to the three-color combinations. Thus, we used the three colors most frequently used in a painting to estimate its mood.

In general, a digital color image has a color depth of $24$ bits. Such an image has too many discrete colors, resulting in attempts to ascertain the most frequently used color being a meaningless exercise. For this reason, we first normalize an image by enforcing a limited number of colors. Kobayashi [3] used the Hue & Tone $130$ system to construct a three-color combination image scale. We utilized the same color system to normalize the colors of the image.

After normalizing the colors, we estimated the coordinates of moods in the color image scale from the image of a painting by using the prediction function acquired in Section II. Kobayashi indicated the coordinates of 16 famous paintings in [3] (Figure 2(a)). To conduct a similar activity for the name of moods, we acquired the coordinates of paintings by calculating the center position of each painting on the figure in [3]. For $16$ paintings with ground truth mood, we estimated each mood as coordinates in the color image scale (Figure 2(b)). In our experiment, the magnitude of the mean error was recorded as $2.08$.

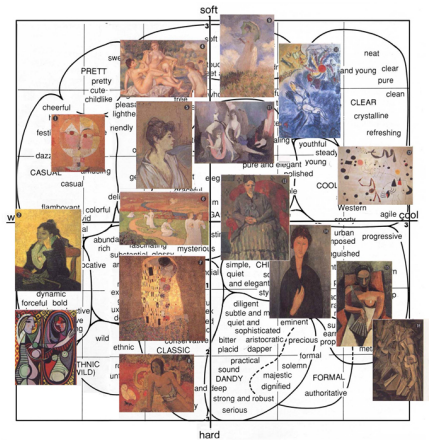
In our experiment, mood estimation performance from paintings was lower than that of mood estimation from three-color combinations. In general, the colors in digital images of the same painting differ slightly, such that prediction depends primarily on the color of the each image in isolation. In our experiments, we did not utilize the same exact images that were employed by Kobayashi; thus, the predicted mood differed from Kobayashis ground truth. Moreover, we obtained our three-color combinations from an image via the naïve approach; hence, there was no guarantee that the extracted three-color combination precisely represented the image. Therefore, a more robust approach for obtaining three-color combinations from an image is required.

## IV. Conclusion and Future Work

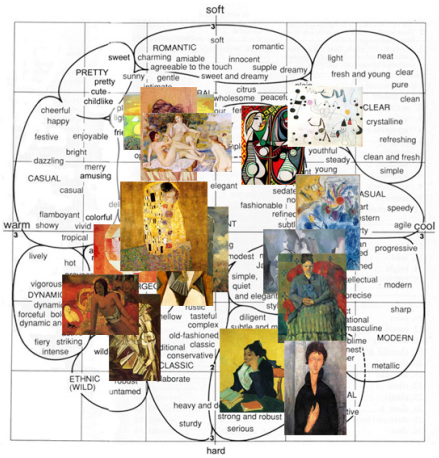
In this paper, we estimated the mood of paintings using the three-color combinations color image scale. First, we defined the features of a three-color combination as several properties of each color. Then, we established a mood prediction function from feature-mood pairs using linear regression. Finally, we extracted three dominant colors from images of paintings, and estimated moods by using the prediction function.

Our experiment depended predominantly on Kobayashis data. However, for the same painting and colors, the mood evoked varies according to the era, age, culture, type of education, and types of color blindness. In future work, we will survey the mood of more paintings and colors in relation to todays society. In addition, we will consider various individual conditions, such as age, education, and cultural area.

In addition to color, several other factors affect the mood of a painting. We plan to determine those other factors and refine our prediction function by incorporating them in our analysis.




(a) Ground truth provided by Kobayashi




(b) our results

Figure 2. Mood estimation results.

## Acknowledgment

This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2014.

## References

[1] D. Joshi, R. Datta, E. Fedorovskaya, Q.-T. Luong, J. Wang, J. Li, and J. Luo, "Aesthetics and emotions in images," Signal Processing Magazine, IEEE, vol. 28, no. 5, Sept 2011, pp. 94–115.

[2] S. Kobayashi, "The aim and method of the color image scale," Color Research & Application, vol. 6, no. 2, 1981, pp. 93–107. [Online]. Available: http://dx.doi.org/10.1002/col.5080060210

[3] ——, Color Image Scale. Kosdansha International, 1991.

[4] G. A. Seber and A. J. Lee, Linear regression analysis. John Wiley & Sons, 2012, vol. 936.

 ISBN: 978-1-61208-392-6

# The New Computational Analysis Tool for Upper Limb Movement in Physiotherapy Biomedical Applications

Jesse Teixeira da Silva
Department of Technological and Professional Education
Federal University of Paraná - UFPR
Curitiba, Paraná State, Brazil
Email: jessesilva@yahoo.com.br

José Elmar Feger
Department of Technological and Professional Education
Federal University of Paraná - UFPR
Curitiba, Paraná State, Brazil
Email: elmar.josefeger@gmail.com

Luiz Antônio Pereira Neves
Department of Technological and Professional Education
Federal University of Paraná - UFPR
Curitiba, Paraná State, Brazil
Email: lapneves@gmail.com

***Abstract*—The present project aims at developing an interface capable of capturing, analyzing and stimulating movements of persons who have some motor dysfunction in the upper arms due to accidents or congenital disabilities, i.e., those who need rehabilitation treatment. In this paper, we have used accelerometers with the objective of capturing the movements carried out by the user. These movements are captured in the form of analogical signals represented by voltage variation. Hence, it requires the use of a microcontroller to handle the data. This microcontroller controls how the data will be sent to the computer, and also, it makes possible the transformation of the signals into the digital format, which facilitates the processing steps and the computer analysis. When the obtained data are sent to a computer, their numerical transformation step starts off with speed and it is plotted on the user screen in an animated form representing the user's movement. The user receives an instantaneous evaluation of his movements throughout the time spent to execute it. The results obtained from the experiments show that this is a promising tool. As it is a prototype, it still needs several improvements. We believe this tool can become a commercial technology with great accessibility and low cost.**

***Keywords–Computational Physiotherapy; Accelerometers; Rehabilitation Treatment; Biomedical Visualization.***

## I. INTRODUCTION

With the growing use of acceleration sensors in technological products, such as video games [1], mobile phones and computers, we believe the development of a system capable of analyzing and stimulating the recovery of patients who suffer from some minor arms dysfunction as a result of accidents or some pre-existing physical handicap is needed. The benefit of this technology would be of great help and of a strong social and intellectual impact. These technologies offer social and personal benefits to the bearers of motor handicaps [2][3][4].

Therefore, the present research proposes the conceptualization and the building of an electronic device capable of rehabilitating upper limbs using current hardware and software technologies.

In the scientific literature, there are studies about the use of accelerometers for physical evaluation of elderly and children [5][6][7][8], involving gait rehabilitation and very little about rehabilitation in the field of physiotherapy, mainly for upper limb movement. Thus, this fact is, our object of work.

Therefore, the main contribution of this research is a new computational analysis tool for upper limb movement in physiotherapy. The proposed system generates an output containing all the movement information in a plain text format, saving it in an embedded database. The containing information are the positions of the selected arm during the task presented as a challenge by the system to the user. This information can be compared with previous database register to validate the conformity of the movement. The system is also able to compare this information with movements from other users and samples, which guarantee that the relation between the number of attempts and the movement accuracy tends to shorten each time the system is used. Another interesting point is that it sets out the accelerometer information and transforms it in "position in space" information, making possible that this same data can be applied in a range of other organizations for comparison and also permits the application to evolve in terms of hardware and software without losing the historical information, being that way, independent of future technologies and computational architecture.

The present article is organized in the following way: Section II addresses other works depicting the application of accelerometers in medical applications, Section III introduces the technology employed. Section IV describes the analysis of results obtained, and finally, Section V presents the conclusion.

## II. STATE OF ART

Some studies on accelerometers in medical applications can be found in the literature which some of them show the use of accelerometers in intervention programs to promote physical activity in children and adolescents [5][6] and include methods for assessing physical activity focused on older adults [7][8]. Others are using accelerometers for mechanic impact analysis in the prostheses of a bilateral lower-limb amputee during the gait [9], aiming to elicit more precise answers.

Michaelsen et al. [10] proposed to use accelerometers to access to identify of changes in movements of subjects with hemiparesis, showing to be an alternative to measure movements.

With respect to these studies, our research is innovative because it presents a rehabilitation tool to the upper limbs, for children and adults, using a playful manner, a video game for better interaction of the patient with his rehabilitation process.

 ISBN: 978-1-61208-392-6 

## III. Proposed Methodology

Because of the need of improving some of the methods applied in physiotherapy and of facilitating the rehabilitation of patients with upper limbs motor handicaps, the present research proposes an interactive and functional limb movement analysis tool. Therefore, the methodology here described uses an exploratory approach of observational and empiric procedures. To this end, the present methodology is organized in four steps, such as shown in Figure 1.

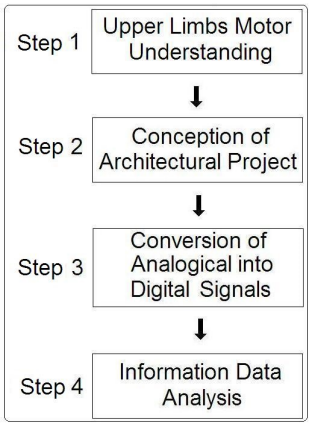



Figure 1. Proposed metnodology.

### A. Step 1 - Upper Limbs Motor Understanding

This step presents an analysis and the drawbacks faced by people with motor handicaps of the upper limbs in order to identify the essential requirements of physiotherapy. This is an observational study.

### B. Step 2 - Conception of Architectural Project

This step consists of definition and conception of proposed architectural project in order to obtain human biometrics data.

### C. Step 3 - Conversion of analogical into digital signals

This step is aimed at the development of a tool capable of converting analogical signals in to digital ones. Digital signals are necessary in order to feed them into the computational tool to be developed.

### D. Step 4 - Information data analysis

This step is aimed at the construction of a computational tool to facilitate analysis and execution of physiotherapy procedures.

## IV. Experiments and Results

### A. Results of Step 1- Understanding of the Upper Limbs Motor Movements

According to the work done by Schonke et al. [11], the human arm is divided into three main parts: the arm, forearm and hand. In a healthy arm, these three parts work together with each movement and no matter how simple the movement is there is an interaction amongst the muscles that cover theses parts of the arm, with each one of them generating a typical rotation, speed and acceleration which can be recognized and standardized by the sampling of mean measurements from different patients. With an unhealthy arm, regardless of the anomaly, the angle formed by the different limbs, the speed of each member, even with a resting position, suffer minimal or marked alterations. The project we developed aims at analyzing these alterations by converting them from the physiologic to the digital form, i.e., if there is any physiologic dysfunction or alteration in this limb, the equipment will be able to recognize and help with the recovery process [12].

### B. Results of Step 2 - Conception of Architectural Project

In order to develop the proposed computational tool, the architecture design project is divided into two parts: Hardware and Software.

*1) Hardware Project:* This project consists of the capture and analysis of signals generated by the accelerometers, besides the transformation of signals from the analogical to the digital form [13]. The circuit here developed counts primarily with the accelerometers [14], which serve as sensors for the capture of movement, as shown in Figures 2 and 3.

Figure 2. MMA7260Q accelerometer coupled in your circuit for ease of manipulation [14].

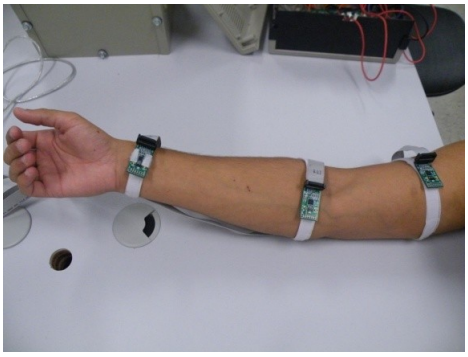


Figure 3. Correct position and fixation of accelerometers on the user's arm.

The accelerometers [13] work as sensors of movement capture. The accelerometers are fixed to the arm in three different positions: wrist, forearm and shoulder. These sensors are capable of analyzing and generating output data in the form of potential differences (volts), and in this way informing whether the region under analysis is at rest or moving, in addition to the movement velocity and direction in three dimensions, all at the same time.

In this phase, data are submitted to a differentiator circuit where the initial potential (offset potential) is compared to that generated by a potential divider circuit and resulting in an output potential difference close to zero.

The three accelerometers are laid on the user arm in the following way: the first lies on the wrist, the second on the

 ISBN: 978-1-61208-392-6 

forearm and the third close to the shoulder and held in position by Velcroed straps (Figure 3).

The three accelerometers capture the movement of the analyzed region during rest or otherwise, the speed and direction of movements, at all times.

After acquisition, data are sent to a differentiator circuit, where the initial potential (offset potential) is close to zero; during the second phase data are sent to a low-pass filter [15], where noise is minimized and input potential is amplified in order to increase the voltage potential band analyzed. The operational amplifiers generate a Chebyshev low-pass circuit with an input voltage potential gain of eight times with a 50Hz cut-off frequency [16]. This voltage potential gain facilitates the analysis and conversion of analogical to digital data given that data will be represented with a maximum of 8 bits. This is illustrated in Tables I and II.

Afterwards, the treated data are sent to a micro-controller, where they are organized and orderly sent to the PC via an USP port [17].

*2) Software Project:* This phase consists in receiving, graphically displaying, analyzing and comparing the digital signal to pre-established movement digital signals with a meaningful concept output to the user.

Therefore, software was developed which is capable to graphically displaying the movement made by the user, who should try to mimic a pre-established movement generated by the system administrator and presented to him.

Firstly, the user visualizes on the computer screen his pre-established handicap motor movement to be performed. After acquisition by the accelerometers, the data are processed.

In the first step, the data are submitted to a differentiator circuit where the initial circuit potential (offset potential) is compared to a potential generated by a potential divider circuit with an output potential close to zero.

In the second step, these potentials are submitted to a low-pass filter with a cut-off frequency of 50Hz to filter out noise and generate eight-fold amplification in order to be analyzed and converted to digital signals by the micro-controller.

In the third step, the generated signal arrives at the micro-controller, where all signals are organized according with each accelerometer and the axis read, and then transformed to the digital format. In the micro-controller, a signal input-output routine was developed using the C programming language.

The newly received information is sent via the USP port to a PC by a USB serial convector cable.

In the final step, the software handles the information. The data are first captured and graphically presented in agreement with the movement performed by the user.

### C. Results of Step 4 - Information Analysis

After the movement has ceased, the software evaluates the values obtained and compares them to the values of pre-established movements stored in the database.

Finally, a concept (comparison with what is already stored in the database) or diagnosis is given to the movement and thereafter the movement may be repeated or the user may perform the next movement.

### D. Experiments

The software developed with JAVA programming language [18] under the NETBEANS development platform [19] controls the input of digitally transformed signals and presents them in a graphic form.

In addition, the software posses a database where the movements to be repeated by the user and the user own movements are stored in order to proportionate a comparative analysis amongst the movements performed during the patient's (user) rehabilitation.

The proposed tool is also capable of comparing the last movement performed by the user to the pre-established movements stored in the database and return a diagnosis (concept). In this way, the user has the option of repeating or advancing to the next movement. Given that all information is stored in the database, the user has the option of recording this information in portable digital media to be taken to other professionals of the medical area.

The testing phase was carried out after the completion of the project implementation phase.

The differentiator circuit output band results obtained during the testing phase are shown in Table I. These results are the average values regarding all the three axis exactly as it is supplied from our accelerometers. In this case, it's possible to see that our zero value would be in 300 millivolts and that way we can notice that always that the value is below 300 mV it means that the arm is moving against the axis and above it towards the axis. For example, if we have the X axis representing the vertical movement and when the arm is moving left, we have a value varying between 100 mV and 300 mV. When the arm is moving right we have this value between 300 mV and 500 mV, depending on the movement velocity.

TABLE I. POTENTIAL VALUES OBTAINED AFTER THE USE OF THE DIFFERENTIATOR CIRCUIT.

| Type/shape of movement | Potential ($mV$) |
|---|---|
| Still arm | $\approx$ 300 |
| Towards the axis | Between 300 and 500 |
| Against the axis | Between 100 and 300 |

In sequence, we used a Chebyshev low-pass filter and then a signal amplifier circuit to improve these signals. With this technique, an eight-fold gain was obtained and the electrical noises were lowered. Our final electric results obtained from the micro-controller input are shown in Table II.

TABLE II. POTENTIAL VALUES OBTAINED FROM THE MICRO-CONTROLLER INPUT WITH THE USE OF A FILTER AND WITH A GAIN OF EIGHT-FOLD.

| Type/shape of movement | Potential ($V$) |
|---|---|
| Still arm | $\approx$ 2,4 |
| Towards the axis | Between 2.4 and 4.0 |
| Against the axis | Between 0.8 and 2.4 |

From the values obtained with the proposed device and displayed in Table II, it is possible to calculate the velocity of the arm movement.

Figure 4 demonstrates the values received from three axes of an accelerometer when strapped to the patients wrist while

 ISBN: 978-1-61208-392-6 

the calibration is running. Note that as expected, the value that represents an arm while not moving is not the same as the electrical value presented in the Table 1 and Table 2. This occurs because the microprocessor delivers us values between 0 and 255, based on its input voltage. That way our software can estimate the velocity and position using a simple conversion routine. Equally, we can understand, the three calibration values obtained after a hundred rounds will be looked at as our zero value.

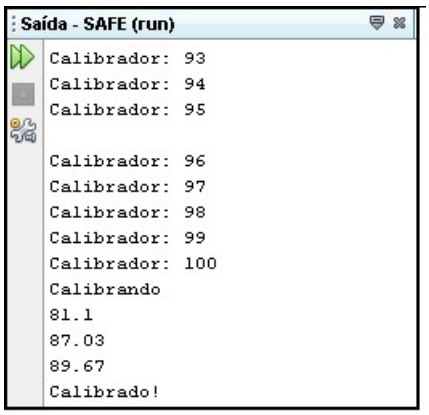



Figure 4. Calibration values from the first accelerometer's three axes, during a test.

Figure 5 illustrates the moment when the user executed a software pre-established movement.


Figure 5. Movement being executed by the user. The movements activate the Mario doll on the path to find his beloved.

The user can repeat the movements as many times as it is necessary and in the same way, the user can see the executed movements during the previous steps. On the screen, the data of the movement carried out by the user can be visualized based on the position related to the axis x and y with the time spent on it (see Figure 6).

In the future work, it will be possible to show the movement in 3D because we already have read the z-axis. However, it will demand one more advanced interface and it will require more powerful equipments.

The result of calculations performed by the software is

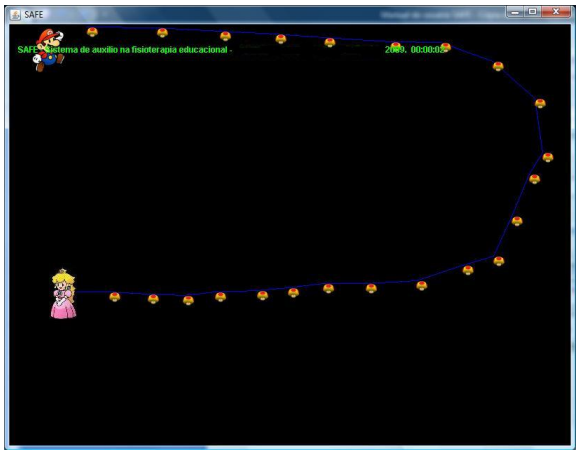



Figure 6. Another example of movement being executed by the user, more simply, Mario to his beloved.

a smooth movement and reflecting precisely the user performance. This makes the use of this tool a viable option for physiotherapy and revealed several improvement opportunities and new applications.

The potential values obtained are translated into a visually comprehensible screen readout, which allows a satisfactory interaction between patient and software output.

## V. Conclusion

After the end of testing and performance proofing of the proposed device, named SAFE, we understand that it is a prototype that still needs further improvement in order to have its use in the physiotherapy field.

The biggest hurdle faced in this project was to obtain knowledge of this new technology and its implementation within the academic environment, i.e. the accelerometers. In spite of their presence as portable devices in the market place, these tools still lacks specialists. This prompted us to search for help and assistance from companies and experts abroad and illustrated how complex and painstaking the search for new solutions for the problems faced within this project are.

When analyzed from the academic point of view, we believe that the project proposal was successfully achieved. This conclusion stems from the ability to communicate and present the data obtained and from the employment of the information obtained. Besides, we believe the results obtained here will stimulate a great spectrum of future opportunities of implementation based on this solution.

Given that the programming language used here is considered slow by many when compared to more traditional languages, we believe we managed to make JAVA programming language an ally in the task of showing the user, in a visual way, the arm movements to be used in physiotherapy treatment.

It is worth mentioning that the results here obtained from the executed tests with the accelerometers should present better results when a newer generation of devices are used. This move will allow a lesser compensation from software to hardware deficiency and allowing a more precise and faster communication and processing of information.

There is no complexity in the hardware used and this allowed our experience to be of an easier nature and helped

 ISBN: 978-1-61208-392-6 

us to elaborate new solutions to the problems faced, in addition to making this project more viable and economically advantageous.

## REFERENCES


[1] P. Cordeiro, 2009, physiotherapy innovatives in treatment with Video Games [online] Available at http://www.acessa.com/saude/arquivo/noticias [retrieved: July, 2014].

[2] J. Hackett, B. Johnson, A. Parkin, and T. Southwood, "Physiotherapy and occupational therapy for juvenile chronic arthritis: custom and practice in five centres in the uk, usa and canada," British Journal of Rheumatology, vol. 35, 1996, pp. 695–699.

[3] S. Lianza, Rehabilitation Medicine. Medicina de Reabilitação, 2002.

[4] J. Suri, D. Wilson, and S. Laxminarayan, Handbook of Biomedical Image Analysis: Volume 3: Registration Models , I. T. in Biomedical Engineering, Ed. Springer-Verlag New York, Inc, 2005.

[5] R. I. Medina-Blanco, A. Jimenez-Cruz, M. E. Perez-Morales, A. L. Armendariz-Anguiano, and M. Bacardi Gascon, "Programas de intervencion para la promocion de actividad fisica en ninos escolares: revision sistematica," Nutricion Hospitalaria, vol. 26, 04 2011, pp. 265 – 270.

[6] M. Romanzini, E. L. Petroski, and F. F. Reichert, "Thresholds of accelerometers to estimate the intensity of physical activity in children and adolescents: A Systematic Review," Brazilian Journal of Kinanthropometry and Human Performance, vol. 14, 00 2012, pp. 101 – 113.

[7] D. T. Ueno, E. Sebastiao, D. I. Corazza, and S. Gobbi, "Methods for assessing physical activity: a systematic review focused on older adults ," Brazilian Journal of Kinanthropometry and Human Performance, vol. 15, 04 2013, pp. 256 – 265.

[8] R. F. I. Osti, L. M. T. Garcia, and A. A. Florindo, "Validation of the 24-hour physical activity recall in elderly adults," Brazilian Journal of Kinanthropometry and Human Performance, vol. 16, 00 2014, pp. 15 – 26.

[9] J. F. Sousa, J. C. Ribeiro, C. C. Sa, A. Novo, and V. P. Lopes, "Mechanic impact analysis in the prostheses of a bilateral lower-limb amputee during the gait," Fisioterapia e Pesquisa, vol. 18, 03 2011, pp. 11 – 16.

[10] S. M. Michaelsen, R. P. Gomes, A. P. Marques, L. C. Rodrigues, N. G. B. Junior, R. Claudino, and M. J. dos Santos, "Using an accelerometer for analyzing a reach-to-grasp movement after stroke," Motriz: Revista de Educacao Fisica, vol. 19, 12 2013, pp. 746 – 752.

[11] M. Schonke, E. Schulte, U. Schumacher, M. Voll, and K. Wesker, Prometheus - Atlas De Anatomia Geral e Aparelho Locomotor. Guanabara Koogan, 2009.

[12] E. L. Sanchez, "History of physical therapy in Brazil and in the World," Brazilian update Physiotherapy Journal, Ano II. Vol. I. N. 03, Mai/Jun., Sao Paulo, 1984.

[13] S. Kurt and C. Oscar, "Implementing positioning algorithms using accelerometers," Freescale Semiconductor, vol. 02, 2007.

[14] 3 Axis Acceleration Sensor Board User's Guide. Sure Electronics Inc, 2007.

[15] A. P. Junior, Analog electronics: operational amplifiers and active filters. Porto Alegre: Bookman, 2003.

[16] F. Najm, "Low-pass filter for computing the transition density in digital circuits," IEEE Transactions on Computer-Aided Design, vol. 13, 1994, pp. 1123–1131.

[17] S. Franco, Design with Operational Amplifiers and Analog Integrated Circuits. New York: Mc-Graw-Hill, 2000.

[18] B. Eckel, Thinking in Java, 4th Edition. Prentice Hall, 2006.

[19] 2009, welcome to NetBeans [online] Available at http://www.netbeans.org/ [retrieved: July, 2014].


 ISBN: 978-1-61208-392-6 

# Towards Open Data for Personal Web Tasking

Anna Ruokonen

Department of Electronic Engineering
City University of Hong Kong
Kowloon Tong, Hong Kong
Email: aruokone@cityu.edu.hk

Otto Hylli and Samuel Lahtinen

Pervasive Computing
Tampere University of Technology
Tampere, Finland
Email: firstname.lastname@tut.fi

***Abstract*—Open data is targeted for public use and usually released by the public sector, e.g., by the government. Its utilization, however, is currently restricted to programmers with knowledge on programming and processing of different data formats. The actual targeted end-users are simply lacking sufficient knowledge and technical skills. End-user tools for data visualization and browsing are available, but customization of the presented data and interaction patterns are limited. The users are unlikely to be able to exploit their creativity and meet their personal needs. In this paper, an on-going work on enabling open data sets for personal Web tasking is presented. Personal Web Tasking aims at automation of user-centric Web interactions. Work presented in this paper concentrates on automation of open data queries defined by the end-users.**

***Keywords–Personal Web Tasking; Open Data Queries; End-User Development.***

## I. Introduction

The Web is becoming increasingly data-centric, consisting of interconnected data-oriented resources, large data sets and semantic linked data. Unlike HyperText Markup Language (HTML), which is designed to format and present text, EXtensible Markup Language (XML) is designed to structure data. With XML, individual data elements can be put in context of a larger taxonomy in order to turn plain text into data. While the main consumers of HTML documents are humans, the main data consumers are computers.

More and more open data sets are made available for the public, especially Public Sector Information (PSI) provided by the governments. Open means anyone can freely access, use, modify, and share for any purpose [1]. Also, by the definition, open data needs to be available in a machine-readable format, like Comma-Separated Values (CSV), XML, or Application Programming Interfaces (API). However, there is no generic mechanism for citizens to explore and consume the data sets [2] [3]. Thus, they depend on software developers to build proper tool-support for data consumption. These applications, however, are often specific to certain predefined data sets and provide limited possibilities for interaction.

End-user development aims at empowering non-programmer end-users to create their own applications. When building systems for end-user computation, the essential thing is to provide an environment, which supports the end-users' concepts and mental models [3]. Especially when consuming data, abstraction and visualization are an essential step to enable end-user interaction [4] [5].

In our previous work, we have studied description of Web resources such that they support end-user driven development [6]. It means that in addition to technical details, the service descriptions should include also end-user targeted metadata, including descriptions of parameters and a type of operations in a way that the concepts are familiar to the end-users.

Personal Web Tasking (PWT) aims at automation of user-centric repetitive web interactions by assisting users in achieving their personal goals on using internet systems [7]. A simple example of a personal Web task utilizing open data is monitoring of Air Quality Health Index (AQHI). It extracts the recent AQHI value and sends the end user a notification, if the value evaluates unhealthy. The PWT system executes the data query repeatedly on behalf of the user, e.g., every hour.

In this paper, an on-going work on end-user approach for creating personal Web tasks involving structured open data sets, provided as XML format or as an API, is presented. Data tasks are based on indicators, which are data elements selected by the end-user for monitoring. The task engine will automatically execute queries on the indicator values. At the current stage, a simple visualization of the data set is provided to enable the end-user identify certain indicators in the data set.

The rest of the paper is organized as follows. Related work is presented in Section II. In Section III, our approach for personal Web tasking with structured open data is presented. Challenges on building such end-user tasking system are discussed in Section IV. Finally, our plans for future work and conclusions are presented Section V.

## II. Related work

There are existing approaches that focus on personal Web tasking, data visualization, or creation of personalized data feeds [5] [8]–[10]. Our approach aims at combining these aspects. Especially, existing approaches of personal Web tasking do not consider utilization of open data.

A self-adaptive application aims at predicting users preferences within a dynamic environment. It involves monitoring the user behavior and changes in the context and system infrastructures. It can be used to enable PWT systems that adapt on changing context and users personal goals to provide pleasant user experiences. In [8], Castaneda *et al.* present a self-adaptive approach where task personalization is done based on context information, such as user's personal context, and historical interactions and social network. Whereas these works emphasizes the user experience, we believe that the power of service-oriented systems and open data can be fully exploited when users are able to develop their own personal Web task consisting of service interactions.

 ISBN: 978-1-61208-392-6 

In [9], Ng J. presents an approach to extend REST for smarter interactions to support personal Web tasking, called REpresentational Action State Transfer (REAST). A key feature related to personal Web tasking is a meta-model, which can be used to build user interface (UI) widgets to compose tasks without programming. REAST interactions are based on transitions from one action state of a resource to another action state of another resource. REAST actions are *Create*, *Read*, *Update*, *Notify*, and *Share*. Notify and Share can be used to enable social networking among users. In addition, event-based and time-based constraints can be attached to customize action execution by the end-users. It is assumed that information technology (IT) personnel first develop a configurable REAST application with supported UI widgets. The end-users can compose their personal tasks using the predefined UI widgets and resources. Our goal is that the user can select the resources and the system automatically generates the UI widget. Furthermore, the approach presented in [9] does not consider interaction with open data.

Castellani *et al.* present a web application framework for Internet of things (WebIoT) [10]. The framework is implemented based on Google Web Toolkit. It utilizes thing-centric design and communication is done using REpresentational State Transfer (REST) paradigm. It aims at integration of services, devices, sensors and other smart things into the Web 2.0 by allowing users to develop and exploit their own applications. Thus, sharing the same vision of user-centered development as emphasized in our work.

Generic mechanisms for data visualization, a tool not specific to any particular data set, for the end-users has been studied in [5]. The work presented by Mazumdar *et al.* aims at identifying the end-users' needs and preferences. Especially, they provide dashboard alternative visualizations, e.g., different kind diagrams, of data content is provided. The users are allowed to use filters to reason on the data. Filters are implemented as SPARQL Protocol and RDF Query Language (SPARQL) [11] queries. In our approach, a simpler approach for data interaction is taken, the user is only assumed to select indicators for future monitoring.

In our earlier work [6], we have presented an approach of adding metadata on the service descriptions. A similar approach, adding annotations on the open data sets, can be applied to provide useful information for the end-users and help them to develop their personal Web tasks. There is also existing work on providing guidelines for visual overviews of open government data [2].

## III. Personal Web tasking with open data

As a part of our earlier work, we have presented tool-support for composing service interactions into personal Web tasks [12]. It is now extended to provide the end-users a simple mechanism to query and interact with open data sets. Sequences of end-user defined Web interactions, like Representational State Transfer (REST) service invocation, data queries are supported as personal Web tasks. Execution of the tasks is handled automatically by a task engine. An overview of the approach is shown in Figure 1.

### A. Data abstraction and visualization

To enable end-user interaction, an abstraction and visualization of the data set must be provided. At the current


Figure 1. Personal Web tasking system

stage of our prototype implementation, a simple approach is implemented. The abstraction of the data set is provided by reducing the document structure. It is implemented as a simple algorithm that parses the XML feed. It finds the elements with the higest occurency and visualizes them to the user. For example, Figure 2 presents an air quality report AQHI24HrReport provided by the Hong Kong government as an XML feed. It contains 360 *item* elements presenting the resent measurement entries and some other XML elements presenting some document level information. After abstraction, only the structure of the *item* element and the latest values are presented to the end-user. Visualization the data content is shown in Figure 3.

```
<AQHI24HrReport>
<title>Environmental Protection Department - AQHI</title>
<link>http://www.aqhi.gov.hk</link>
<description>Environmental Protection Department -
   Past 24-hr AQHI</description>
<language>en-us</language>
<copyright>Environmental Protection Department</copyright>
<webMaster>enquiry@epd.gov.hk</webMaster>
<lastBuildDate>Fri, 26 Sep 2014 18:30:00 +0800</lastBuildDate>

<item>
   <type>Roadside Stations</type>
   <StationName>Mong Kok</StationName>
   <DateTime>Fri, 26 Sep 2014 18:00:00 +0800</DateTime>
   <aqhi>5</aqhi>
</item>

..
</AQHI24HrReport>
```

Figure 2. AQHI24HrReport

The example in Figure 3 shows the creation of a data task. First, the user selects a data set by defining the URL or selecting one of the data sets stored in the PWT system. The system extracts the data sets and presents an abstraction of its' contents to the user as a table. The users can select which parts of the data they want to monitor. Selected elements are included in the data task as data indicators. In the presented scenario, the user wants to connect to her Twitter account and send an automated tweet, if the air quality is unhealthy. Thus, she defines an interaction sequence, connecting the result of the data query, *if* control node, and Twitter service.

Figure 3 shows the creation of a simple control structure

 ISBN: 978-1-61208-392-6

with a condition. The user can select the desired operator from the operator list and specify the operand value. For numeric values <, >, and *equals* operators can be used. For strings, *equals*, *contains substring*, *prefix*, and *postfix* operators can be used. On the first column the user can select which indicators she wants to be included in the constraint. The first column presents the name of the data element. The third column shows supported operators, which can be used to create constraints on the indicator value. *ANY* means no constraint is used. The desired constraint value can be specified on the fourth column. The default values are the latest data values. In the given example, the user wants to monitor air quality of the Central/Western district of Hong Kong. If the air quality reaches unhealthy (*aqhi* > *6*), the user wants to send a tweet.

Before completing the task and sending it to the task engine, the user specifies the task execution frequency. For example, if execution frequency is specified to be one hour, the data query is automatically executed every hour by the task engine.

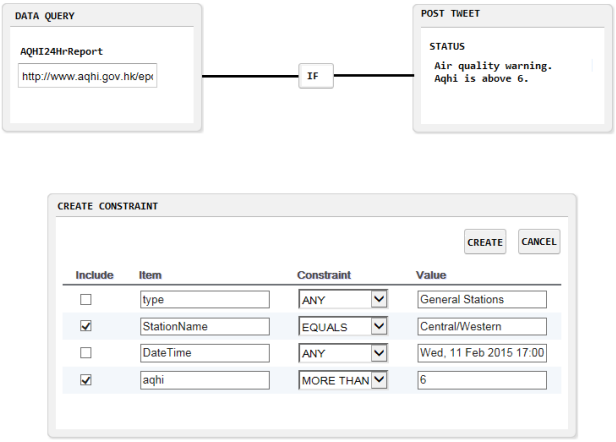



Figure 3. Designing a data task

## B. Task creation with a data set

In the proposed Web Tasking system, tasks are composed of a sequence of activities. Basic activities are the CRUD operations to be used to interact with RESTful services: *POST*, *GET*, *PUT*, and *DELETE*. In addition, there are two activies *getDATA* and *NOTIFY* activity. *getDATA* activity can be attached to a data feed to simply get the latest value of a certain indicator or data feed. A *NOTIFY* activity is used to create a custom user notification. User defined constraints are implemented as variables and *if* statements in our task descriptions. Each activity takes a set of input parameters, which are required for its execution. For example, *getDATA* activity has two input parameters, the data URL and an indicator identifier. The values are extracted from the end-user input.

For example, as a result of scenario presented in Figure 3, an instance of *getDATA* activity is created. The parameters are the data *url* http://www.aqhi.gov.hk/epd/ddata/html/out/24aqhi_Eng.xml, and indicator identifiers *StationName* and *aqhi*. The output of the *getDATA* activity is connected to a *if* control node with the created condition (*StationName* equals *CentralWestern* and *aqhi* > 6). If the condition evaluates true, the task execution proceeds to invoking Twitter API to send a status update with a warning message. This is done with Twitter's *POST* method. *status* parameter defines the warning message. When a new task is created, the user specifies global parameter related to task execution, related to timing and repetition.

# IV. Requirements for personal Web tasking system

The most important components of the PWT system are the task editor and the task engine. The overall architecture including the core components is shown in Figure 4. The task editor is a client side application, which provides a graphical interface for task creation. It can be used with a standard Web browser and it is targeted for the end-users. When the user has finished, the task editor generates an exportable task description given in XML-based language. The task description is stored and executed by the task engine. The task engine also provides a registry for services and data sets. The metadata editor, shown in Figure 4, is designed for augmenting service descriptions.

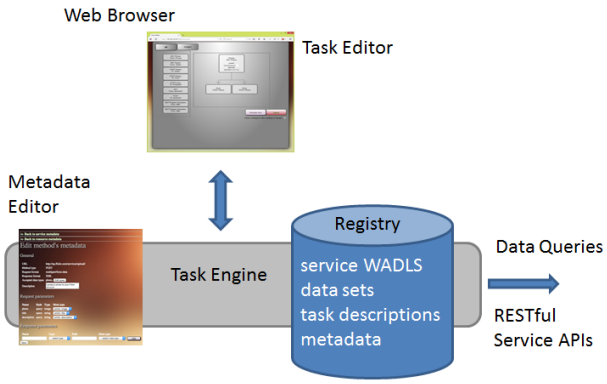



Figure 4. High level architecture

The overall personal Web tasking system must support the following activities: (1) task design, (2) Web interaction sequencing, and (3) task execution.

## A. Task design

The main requirement for our end-user driven approach is to enable the user interaction with the data sets and RESTful APIs. Support for data interaction is provided thought data abstraction and data visualization as presented in the example in Section III. In the presented approach, only structured XML documents are considered as the data format. To provide a simple mechanism to interact with the data, our approach is based on monitoring single values, so called indicators. The selected indicators can be attached with simple constraints. For evaluating the indicator values, basic XML processing tools are used.

In our previous work, we have presented an approach for adding meta-data in RESTful service descriptions to support end-user driven service configurations. It assumes no modification on the existing services, but it requires, however, pre-registration of available service descriptions given in Web Application Description Language (WADL) [13]. A similar approach by annotating XML-based data sets can be used. However, this study presents a generic approach with no

 ISBN: 978-1-61208-392-6 

advance registration of the annotated data sets. Thus, the use of metadata is limited to the XML tag names and the XML document structure.

### B. Web interaction sequencing

Sequencing of Web interaction, such as data queries and RESTful service invocation, is supported by an XML-based task description language. An example of sequencing with the task description language is shown in Figure 5. It describes the interaction sequence of the air quality monitor task, containing two activities, the data query (*getDATA* activity) and posting a tweet on Twitter (*POST* activity).

```
<sequence>
   <getDATA url="http://www.aqhi.gov.hk/epd/
   ddata/html/out/24aqhi_Eng.xml" indicator=
   "/AQHI24HrReport/item[StationName='CentralWestern']
   [last()]/aqhi/text()">
   aqhi</getDATA>
         <if variable="aqhi" operator="bigger_than" value="6">
               <POST>
                  <request>status</request>
                  <response/>
                  <resource_id>twitter</resource_id>
               </POST>
         </if>
</sequence>
```

Figure 5. Task description

Variable *aqhi* is used to store the air quality index extracted from the data set and *status* contains the message that will be posted on Twitter. *getDATA* element defines the data set URL and how to get the indicator value. The *if* element takes a variable (the air quality), an operator (bigger than) and a value (6) and evaluates the expression (aqhi >6). If the condition evaluates true the *POST* activity is executed and and the message is posted to Twitter.

## V. Conclusions and future work

More and more open data is made available for the public, especially in the public sector (e.g., European Union and PSI Directive). However, there is no generic mechanism for citizens to explore and consume the data. The users should be provided a simple mechanism to query and interact with the data sets. Our aim is to develop a personal Web tasking system, which supports the use of open XML data composed of a sequence of Web interactions. We present a generic approach, which automatically generates data abstraction and visualization of the data content enabling the end users to interact with the data set. The idea of personal Web tasking is to automatize repetitive Web interactions. For example, do a data query on the data source every hour to get the latest value of a certain indicator the user is interested in. To customize the data task, the user can define simple control structures with constraints. The end-user can, e.g., choose to receive notifications based on the latest data values.

The presented approach will be integrated with our previous work, which provides a platform for personal Web tasking with RESTful services. Interaction with the RESTful services is implemented via augmented WADL descriptions. RESTful service invocations and data queries along with simple control structures can be used as constituents of personal Web tasks, which are executed by the task engine. The ultimate goal of the presented approach is to empower end-users to create their own small Web applications and to benefit from the Internet of resources available. Especially, we want to enable non-programmers to explore and benefit from available public sector data.

## References


[1] The Open Definition, Open Knowledge Source Code, http://opendefinition.org/od/ [accessed: 2015-01-02].

[2] A. Graves and J. Bustos-Jimnez, "Towards visual overviews for open government data," in DataWiz2014: Data Visualisation Workshop, ACM Hypertext 2014 conference, September 2014, pp. 1–6.

[3] J. Rode, M. Rosson, and M. P. Quiñones, "End user development of web applications," in End User Development, ser. Human-Computer Interaction Series, H. Lieberman, F. Patern, and V. Wulf, Eds. Springer Netherlands, 2006, vol. 9, pp. 161–182.

[4] A. Dadzie and M. Rowe, "Approaches to visualising linked data: A survey," Semantic Web, vol. 2, no. 2, Apr. 2011, pp. 89–124.

[5] S. Mazumdar, D. Petrelli, and F. Ciravegna, "Exploring user and system requirements of linked data visualization through a visual dashboard approach," Semantic Web, vol. 5, no. 3, 2012, pp. 203–220.

[6] O. Hylli, S. Lahtinen, A. Ruokonen, , and K. Systä, "Resource description for end-user driven service compositions," in IEEE Services 2014, 2nd International Workshop on Personalized Web Tasking (PWT 2014), July 2014, pp. 11–17.

[7] L. Castaneda, H. Muller, and N. Villegas, "Towards personalized web-tasking: Task simplification challenges," in 2013 IEEE Ninth World Congress on Services (SERVICES), June 2013, pp. 147–153.

[8] L. Castañeda, N. M. Villegas, and H. A. Müller, "Self-adaptive applications: On the development of personalized web-tasking systems," in Proceedings of the 9th International Symposium on Software Engineering for Adaptive and Self-Managing Systems, ser. SEAMS 2014. New York, NY, USA: ACM, 2014, pp. 49–54. [Online]. Available: http://doi.acm.org/10.1145/2593929.2593942/ [accessed:2015-01-02]

[9] J. W. Ng, "Adapting rest to reast building smarter interactions for personal web tasking," in IEEE Services 2014, 2nd International Workshop on Personalized Web Tasking (PWT 2014), July 2014, pp. 38–46.

[10] A. Castellani, M. Dissegna, N. Bui, and M. Zorzi, "Webiot: A web application framework for the internet of things," in Wireless Communications and Networking Conference Workshops (WCNCW), 2012 IEEE, April 2012, pp. 202–207.

[11] SPARQL Query Language for RDF, W3C, http://www.w3.org/TR/rdf-sparql-query/ [accessed: 2015-01-02].

[12] O. Hylli, S. Lahtinen, A. Ruokonen, and K. Systä, "Service composition for end-users," Acta Cybernetica, vol. 21, August 2014, pp. 383–399.

[13] Web Application Description Language (WADL), W3C, http://www.w3.org/Submission/wadl/ [accessed: 2015-01-02], 2009.


 ISBN: 978-1-61208-392-6 

# Analysis and Evaluation of UI Component Integration for Large-Scale Web Content Reuse

Hao Han* and Yinxing Xue†
*Department of Information Science, Faculty of Science
Kanagawa University, Japan
han@kanagawa-u.ac.jp
†Temasek Laboratories
National University of Singapore, Singapore
tslxuey@nus.edu.sg

***Abstract*—Mashup has promoted new creativity and functionality to Web applications through the integration of information/knowledge from multiple websites. However, there is no uniform interface to access the data, computations and user interface provided by the different kinds of Web content. Without open APIs, it is difficult to integrate Web applications with other Web content. In this paper, we present a framework for flexible and lightweight integration of Web content for personal use. We propose a simple Extensible Markup Language (XML) based Web content description language to define the Web content and configure the mashup applications. We also conduct an exploratory analysis and evaluation of user interface components for large-scale reuse.**

***Keywords–UI Component; API; Integration; Retrieval; Security.***

## I. INTRODUCTION

The availability of information/knowledge is increasing explosively on the Web with the development of the Internet, however useful information is not always in a form that supports end-user requirements. This is because there are no uniform interfaces to access the data/computations (application logic) or user interfaces provided by the different kinds of Web content. For reuse and integration, the content used in the majority of the current mashup applications is typically obtained from third party sources through public Web service Application Programming Interfaces (API). The integration of general Web applications requires additional efforts including programming and configuration compared to single-type Web service integration. Such integration is beyond the skills of typical users and restricted to specific technologies or domains. In this paper, we present a description-based framework of flexible and lightweight integration of Web content for customized reuse. We define Web Content Description Language (WCDL) to configure the mashup Web applications. Furthermore, we conduct an analysis and evaluation of User Interface (UI) components for further large-scale reuse.

The remainder of the paper is organized as follows. In Section II, we present the motivation for our study and an overview of the related work. In Section III, we explain description-based integration. We analyse and evaluate the proposed approach for large-scale reuse in Section IV. Finally, we conclude our approach and discuss future work in Section V.

## II. MOTIVATION AND RELATED WORK

The Majority of Web mashup technologies are based on a combination of Web services and Web feeds. Yahoo! Pipes [1] is a composition tool to aggregate, manipulate, and mashup Web services or Web feeds from different websites with a graphical user interface. Mixup [2] is a development and runtime environment for UI integration [3]. It can quickly build complex user interfaces for easy integration using available Web service APIs. Mashup Feeds [4] and WMSL [5] support integrated Web services as continuous queries. They create new services by connecting Web services using join, select, and mapping operations. Similar to these methods, other service-based methods [6], [7], [8] are limited to the combination of existing Web services, Web feeds, or generated Web components.

For the integration of parts of Web applications without Web service APIs, partial webpage clipping methods such as C3W [9] and Marmite [10] are widely used. Users clip a selected part of a webpage, and paste it into a customized webpage. Marmite [10], implemented as a Firefox plug-in using JavaScript and XUL (XML User Interface Language [11]), employs a basic screen-scraping operator to extract the content from webpages and integrate it with other data sources. The operator uses a simple XPath pattern matcher and the data is processed in a manner similar to UNIX pipes. Intel MashMaker [12] is a tool for editing, querying, manipulating and visualizing continuously updated semi-structured data. It allows users to create their own mashups based on data and queries produced by other users and remote sites. However, these methods can only extract Web content from static HyperText Markup Language (HTML) pages. Moreover, it is not easy to realize the interaction between different Web applications such as Safari Web Clip Widgets [13], unless there is a special Web browser and plug-in.

The majority of the current methods are based on existing Web service APIs or Web feeds, require a professional Web programming ability [14], [15], or have other limitations (e.g., the components must be produced by portlets [16] or WSRP [17]). To address these issues, we propose a flexible and lightweight description-based mashup of Web content. Our integration framework facilitates the integration of various Web applications and services for typical users with minimal programming experience. The range of content is extended from Web services to the "traditional Web" (data, computations,

 ISBN: 978-1-61208-392-6 

and the UI components of general Web applications) including content dynamically generated by client-side scripts or plug-ins.

## III. DESCRIPTION-BASED INTEGRATION

The propose of framework is based on the description and integration of Web content, which includes Web services, Web feeds and Web applications. As illustrated in Figure 1, we begin by describing the target Web content in a WCDL file. Then, client requests are sent to the target websites. According to the defined WCDL file, partial information is extracted from the responding webpages if the target Web content derives from Web applications. The XML-based content (returned from the Web services or extracted from the responding webpages) is transformed into an HTML format and the target Web content is integrated.

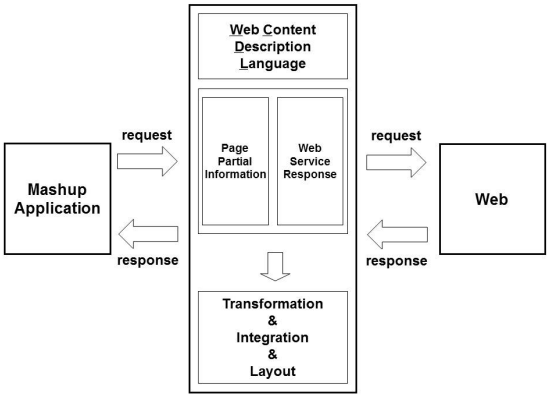



Figure 1. Outline of the proposed description and integration approach/system

"WCDL"is XML-based and intelligible to typical users. In this paper, a typical user is specified as a user who has basic computer knowledge and operation skills without professional programming ability/experience. Compared with standard Web services, it is not easy to access and integrate Web applications because web applications are designed for browsing, not for parsing by program. Without interfaces such as Simple Object Access Protocol (SOAP) [18] or Representational State Transfer (REST) [19], an extraction is used to acquire the target content, and emulation is utilized to realize the automated process of sending requests and receiving responses.

As indicated in Figure 2, WCDL contains the following items to describe the necessary Web application information reflecting end-user operations (e.g., users find the text input field, input the keywords, submit the request, and search for the target content in the response page). *StartPage* is the Uniform Resource Locator (URL) of the webpage of the target Web application where the request of the end-user is submitted. *InputType* is the type of request-input element such as InputBox (text input field), OptionList (drop-down option list), or LinkList (anchor list) in the StartPage. *InputArea* is used to locate the request-input element in the StartPage. If there are other elements with the same InputType in the StartPage, we must define the InputArea.

*ContentType* is the type of target content: static or dynamic. Static content is the unchangeable portion displayed on a webpage after the page is fully loaded and during the viewing process. It contains two kinds of information: property and structure. The property is a text (character string), image (instance of graph), link (hyperlink), or object (instance of video or other multimedia file). The structure is a single, list, or table. Dynamic content is the portion dynamically generated or changed by client-side scripts or plug-ins in a dynamic HTML page according to the users' operations. *ContentArea* is used to locate and extract the target content in the response webpage using an XPath expression. *ContentStyle* is the layout of the target content in the integrated resulting webpage. It is typically limited to static Web content. The extraction results are in an XML format and the style refers to the XML Stylesheet Language Transformation (XSLT) files defined by the end-user. If the webpage containing the target content is a fixed page (e.g., a static webpage with an unchangeable URL), the values of *InputType* and *InputArea* are null (*StartPage* is the target page).

```
<target type="application" ID="...">
    <StartPage>
      URL of webpage where request is submitted
    </StartPage>
    <InputType>
      Type of request-input element in StartPage
    </InputType>
    <InputArea>
      Path of request-input element in HTML document
      of StartPage
    </InputArea>
    <ContentType>
        <type>
          Type of a target part
        </type>
    </ContentType>
    <ContentArea>
        <content>
          Path of a target part in HTML document of
          response webpage
        </content>
    </ContentArea>
    <ContentStyle>
        <style>
          Layout of a target part in resulting webpage
        </style>
    </ContentStyle>
</target>
```

Figure 2. Description of a Web application

Web services and feeds are self-contained and self-describing, and communicate using open protocols such as HTTP. The most widely used style architectures of Web services are SOAP and REST. Recently, new Web services are implemented using the REST architecture rather than SOAP because REST is lightweight, easy to build, and provides readable results. Furthermore, Web feeds such as Atoms or RDF Site Summary (RSS), are usually considered a kind of simplified REST service.

As illustrated in Figure 3, we use the following items to describe REST Web services in WCDL. *BaseURL* is a reference to a method of the target Web service. It contains the hostname, service name, version number, and method name. *Key* is necessary in some Web services. *Query* is the query string of the request URL. The value of each <parameter> is the name of a query parameter. The actual query parameters follow the method (start with a question mark) and assume the form *parameter=value*, where the query parameters and values are URL encoded. Multiple query parameters are separated by an ampersand. The *Type* of a method specifies the format to send requests to the target Web service. The majority of REST API requests use the "GET" method. For the "POST" method,

 ISBN: 978-1-61208-392-6

there is no question mark following the method and the query parameters are passed in the "POST" data block.

*ContentStyle* is the layout of the response content in the integrated resulting webpage. It refers to a layout (template) file. For websites that provide both REST and SOAP Web services, we select the REST Web services as our default selection. For SOAP Web services, we transform these into REST queries (see Yahoo Query Language [20]). The transformation is a semi-automatic process and requires some manual configuration based on the description of the SOAP services in the WSDL [21] files. Once the specification of the target server endpoint URL, namespace declarations, and body elements (Uniform Resource Identifier (URI) of the target object, method name, parameter names) of the SOAP envelope are retrieved, a REST query is generated automatically in the proposed system.

```
<target type="service" ID="...">
    <BaseURL>
      Reference to a service and method
    </BaseURL>
    <Key>
      API key of Web service
    </Key>
    <Query>
        <parameter>
          Name of query parameter
        </parameter>
        ...
    </Query>
    <Type>
      GET or POST method
    </Type>
    <ContentStyle>
      Layout of response content in resulting webpage
    </ContentStyle>
</target>
```

Figure 3. Description of a REST Web service

Based on the abovementioned description in the WCDL file, we search for the target Web content from the Web application. There are two steps during this process. First, we consider the response webpages as the target webpages. Then, we search for the target parts in the response webpages.

Web applications normally provide the request-submit functions for the end-users. For example, search engine applications provide the text input field in the webpage for keyword inputting. Users input the query keywords, submit the request to the server side, and receive the response webpages. To submit a request, there are generally different methods such as "POST" and "GET". For a "more secure" mode, some websites utilize encrypted codes or randomly generated codes during request submitting, which are side-effects of information security. JavaScript can be run as part of submission for validating a request. Hidden objects, which represent hidden input fields in an HTML form, are also widely used to record different submitting options. Consequently, these diverse processing methods can make it difficult to manually parse the HTML or URL template to acquire the target webpages. To obtain the response webpages from all types of Web applications automatically, we implement HtmlUnit [22] to emulate the submitting operation, instead of using the URL templating mechanism. The emulation is based on the event trigger of the element of *InputType* within the *InputArea* of *StartPage*.

*ContentArea* is used to determine the target parts from the webpage. In the tree structure of an HTML document, each path represents the root node of a subtree and each subtree represents a part of the webpage. Response webpages generally have the same or similar layouts if the requests are sent to the same request-submit function. During node searching, if a node cannot be found using a path, a similar path can be tentatively used to search for the node [23].

Once the target parts are located, the content is extracted from the nodes in text format, except for the tags of the HTML document, based on the corresponding *ContentType*. The extracted static content is in an XML format and can be transformed into an HTML document by *ContentStyle*. For dynamic content, we use an effective Hide-and-Display method to control visibility, as the static content extraction method cannot retain the functionality of client side scripts. Currently, many websites use Dynamic HTML (DHTML) or Document Object Model (DOM) [24] scripting technology to create interactive webpages. These contain HTML, client-side scripts, and Document Object Model. The scripts change variables (including the elements outside the target parts such as the hidden values) programmatically in an HTML document and affect the look and function of the static content. If we remove the other parts from the target parts using a traditional static extraction method, the original execution environment of scripts would be broken and the scripts could execute abnormally. Thus, we retain all the parts of each webpage and change the visibility according to the following steps to display the wanted target parts only and hide the other parts of the webpage by setting the property "display" of attribute "style" of all the nodes to "none" (*style.display="none"*).

Except for the dynamic content generated by client-side scripts, both the Web service response and the extracted partial information from Web application are in an XML format. We use an XSLT template processor (*ContentStyle* in WCDL) to transform the XML data into HTML or XHTML documents. After the description, extraction, and transformation, we integrate the content (in an HTML format) from different Web applications or services into a resulting page. We use iframe (or div, span) tags as the default Web content container. This is an HTML element that makes it possible to embed an HTML document into another HTML document. Whereas a regular frame is typically used to logically subdivide the Web content of one webpage, an iframe is more commonly used to insert Web content from another website into the current webpage. Moreover, iframe is supported by all mainstream browsers.

We illustrate an example of the generation of integrable content using a BBC news search. As illustrated in Figure 4, after the user inputs the search keywords and sends the request, the proposed system sends the request to each target website and receives the response content, which is presented as an integrable content segment.

We also generated an example of a mashup Web application. It integrates parts from the following seven Web applications/services/feeds and implements a country-information search function. Table I provides a description of each target Web content. Target A (dynamic contents, mouse move event trigger, client JavaScript): real-time local time; Target B (mashup application): weather information from a mashup application integrated by a weather service and Google Maps service. The weather information part is created by client-side scripts that can respond to click and span events; Target C (static text and image): country's location, basic information and leader's photo; Target D (search engine application): latest corresponding news articles; Target E (flash interaction with

 ISBN: 978-1-61208-392-6 

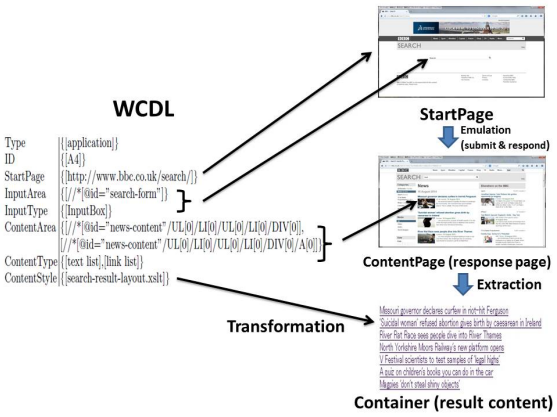



Figure 4. WCDL describes the flow of generation of integrable content

JavaScript): photos displayed with a map that can respond to click events and present the relevant pictures; Target F (Web service): simple introduction to the country; and Target G (RSS/Atom feed): videos about the country. As illustrated in Figure 5, after the user inputs the country name and sends the request, the mashup Web application sends a request to each target website and receives the response Web contents, which are presented in a resulting integrated webpage after layout rearrangement.

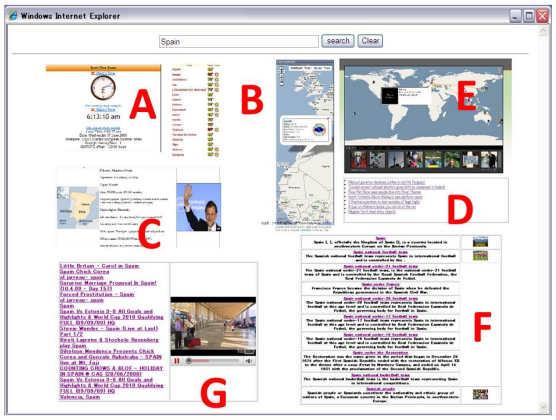



Figure 5. Resulting webpage of an example mashup application

Compared to the single mixing of static webpage segments, in the proposed system, the integrated segments retain their original active functionalities and can be customized into a uniform layout style in a resulting webpage.

## IV. Analysis and Evaluation

We developed a lightweight integration of Web content with simple description and easy configuration based on WCDL. WSDL and WADL [25] are used to describe a series of Web services with details such as abstract types, abstract interfaces, and concrete binding. It cannot be expected that all the tasks will be completed automatically using WSDL or WADL files. For example, the main resource of a WADL description for the Yahoo! News Search application [26] contains more than 40 parameter/attribute values. Compared with the *programmer-oriented* WSDL or WADL, our WCDL provides a shorter and simpler description format, and is applicable to the description of general Web applications. It is easier to read, write, reuse, and update any part of a mashup applications than to work with end-user programming methods. The proposed approach allows a users with no or minimal programming experience to integrate the Web content from multiple Web applications and services. Because the proposed approach is generally applicable to Web applications and services, any content from any kind of Web application becomes integrable

TABLE I. DESCRIPTION OF TARGET WEB CONTENT

| Target | Item | Value |
|---|---|---|
| A | Type | {[application]} |
| | ID | {[A1]} |
| | StartPage | {[http://localtimes.info/]} |
| | InputArea | {[null]} |
| | InputType | {[LinkList]} |
| | ContentType | {[dynamic]} |
| | ContentArea | {[/HTML/BODY/DIV[0]/DIV[2]/DIV[1]/DIV[0]]} |
| | ContentStyle | {[null]} |
| B | Type | {[application]} |
| | ID | {[A2]} |
| | StartPage | {[http://www.weatherbonk.com]} |
| | InputArea | {[//*[@id="searchForm"]]} |
| | InputType | {[InputBox]} |
| | ContentArea | {[//*[@id="grid"],[//*[@id="bonkMapColumn"]} |
| | ContentType | {[dynamic],[dynamic]} |
| | ContentStyle | {[null]} |
| C | Type | {[application]} |
| | ID | {[A3]} |
| | StartPage | {[http://news.bbc.co.uk/1/hi/country_profiles/default.stm]} |
| | InputArea | {[null]} |
| | InputType | {[OptionList]} |
| | ContentArea | {[/HTML/BODY/DIV[0]/DIV[5]/TABLE[0]/TR[0]/TD[1]/ TABLE[2]/TR[1]/TD[0]/TABLE[1]/TR[0]/TD[0]/DIV[0]/ IMG[0]], [//*[@id="content"/UL[0]], [/HTML/BODY/DIV[0]/DIV[5]/TABLE[0]/TR[0]/TD[1]/ TABLE[2]/TR[1]/TD[0]/P[28]/TABLE[0]/TR[0]/TD[0]/ DIV[0]/IMG[0]]} |
| | ContentType | {[image],[text list],[image]} |
| | ContentStyle | {[automatch.xslt]} |
| D | Type | {[application]} |
| | ID | {[A4]} |
| | StartPage | {[http://www.bbc.co.uk/search/]} |
| | InputArea | {[//*[@id="search-form"]} |
| | InputType | {[InputBox]} |
| | ContentArea | {[//*[@id="news-content"/UL[0]/LI[2]/UL[0]/LI[0]/ DIV[0]], [//*[@id="news-content"/UL[0]/LI[2]/UL[0]/LI[0]/ DIV[0]/ A[0]]} |
| | ContentType | {[text list],[link list]} |
| | ContentStyle | {[search-result-layout.xslt]} |
| E | Type | {[application]} |
| | ID | {[A5]} |
| | StartPage | {[http://www.trippermap.com]} |
| | InputArea | {[null]} |
| | InputType | {[null]} |
| | ContentArea | {[//*[@id="maptabs"]} |
| | ContentType | {[dynamic]} |
| | ContentStyle | {[null]} |
| F | Type | {[service]} |
| | ID | {[W6]} |
| | BaseURL | {[http://en.wikipedia.org/w/api.php]} |
| | Key | {[null]} |
| | Query | {[action=opensearch],[format=xml],[search]} |
| | Type | {[GET]} |
| | ContentStyle | {[wiki-layout.xslt]} |
| G | Type | {[service]} |
| | ID | {[W7]} |
| | BaseURL | {[http://gdata.youtube.com/feeds/base/videos]} |
| | Key | {[null]} |
| | Query | {[q]} |
| | Type | {[GET]} |
| | ContentStyle | {[youtube-layout.xslt]} |

and reusable.

Currently, the proposed extraction method is based on the fact that response webpages returned from the same Web application use the same or a similar layout. If the response webpages use different layouts, the extraction precision would degrade because the paths of the target parts vary with the layouts of the webpages (XPaths are changed). Moreover, if the layout of the webpage is updated, users must change the value of *ContentArea* in the corresponding WCDL file (they can use tools such as Firebug [**?**] and are not required to read the source code of the HTML documents). In the proposed integration system, we use emulation to submit requests and acquire responses automatically without manual analysis of typical URLs for templates. Nevertheless, the emulation process of HtmlUnit is slow for some websites and requires more time than the URL templating mechanism if the websites rely on external scripts in the submitting process. The proposed approach is not yet fully automated and may not be efficient in emulation mode.

For the large-scale reuse of UI components, a challenge is to include efficient retrieval of the UI components to facilitate client-side integration. Users currently must locate functionality from websites using general search engines such as Google and Bing, or a library of predefined UI components that is also manually created and limited to a small scale. This conventional method of Web functionality retrieval is inefficient because current general Web search engines primarily use a content-oriented (non-functionality-oriented) search mechanism and information about functionality is beyond the searching scope of these current mainstream search engines. Moreover, this method leads to time-consuming manual verification and comparison if users wish to locate the most desirable and suitable component. In our past experiments and implementations, we were required to spend more than one hour to find and select each suitable component on average, although generating the WCDL for a segment usually required less than 5 minutes (15 senior students of CS were requested to independently complete the construction of mashup application integrating diverse content. The longest time consumed for the WCDL of a segment was 25 minutes and the average time was approximately 4 minutes).

To the best of our knowledge, there is no previous work on large-scale UI component retrieval. We designed a novel searching mechanism [27] that is implemented in the UI component retrieval system. A crawler is used to collect webpages. For webpages containing components, background information such as URL, title, and meta data is collected. For each submitting form and component, diverse attributes (e.g., ID, name, and XPath) are extracted as necessary information and other related information (e.g., alt, hidden values, and defined event trigger JavaScript functions) are extracted as optional information. This information is analyzed and classified to generate a recognition pattern for each Web UI and the index information is then created for component retrieval.

Table II displays a list of component properties as indexes of the BBC news search. *Category*: group of websites/webpages with similar subjects, which are extracted from the meta data of webpage; *Function type*: functionality of UI component; *UI*: UI component type (e.g., text input field, drop-down option list, anchor list); *Request candidate number*: input candidates are unlimited for a text input field or the number of options in an option/link list.“unlimited” is the best applicability of the interface and a larger number represents a better applicability. *Data volume*: size of data transferred between client and component/target website; *Speed*: average time (ms) for the request-response process; *Stability*: statistics of access success (whether the UI component normally runs) rate compiled periodically (100% implies it normally runs during webpage crawling); *Language*: primary language of the website; and *Script*: whether client-side JavaScript runs in the request-response process.

Each index has its unique characteristics. For a more objective evaluation of components, a supervised machine learning approach is employed to perform a construction of ranking models for the further recommendation mechanism in the proposed approach. Features are assigned to represent numerical weightings of indexes such as speed and stability, and manual assessment results are used as training/testing data. Clustering is used to partition components into groups (clusters). The components in the same group are more similar to each other than to those in other groups in *Category*. Users search for the components by providing the component type and other search keywords. Compared with the general content-oriented search engines, this component-oriented search mechanism can achieve faster and more effective component retrieval (within minutes).

TABLE II. PROPERTIES OF UI COMPONENT

| | |
|---|---|
| Category | Application, BBC, Search, news |
| Function type | search |
| UI | text input field |
| Request candidate number | unlimited |
| Data volume | 82.94 KB |
| Speed | 2,759 ms |
| Stability | 100% |
| Language | English |
| Script | no |

However, actually, compared to non-context Web APIs, context-related UI components are coarse-grained. The property information extracted from meta data and tag names is not enough for precise analysis. Moreover, it is also difficult to measure the relative importance of a UI component like PageRank since the main functionality of a UI component is limited in the inner side of a website.

For large-scale reuse of the existing content from Web services or even general Web applications, we must evaluate the scalability of the proposed approach as well as the reusability of the extracted content. On average, it requires approximately 100-1000 ms to extract the contents from a given webpage; there is not a significant difference between a static webpage and a dynamically generated webpage. In the proposed framework, the extraction and the integration are implemented in different components, which allows both steps to be executed in parallel. The extraction step can prebuild a database of possibly useful content that is extracted during crawling. The proposed approach is efficient in both steps and feasible for large-scale reuse, e.g., searching and integrating content from hundreds of thousands of webpages.

Regarding the reusability of the extracted content, the proposed approach realizes the separation of the data and the presentation. Using an XSLT template processor, we can transform the XML-format content into XHTML according

 ISBN: 978-1-61208-392-6 

to the desired layout in the XSLT file. In addition to the data, the extracted content may include UI components. The reusability of these UI components is not undermined as the proposed approach adopts an effective Hide-and-Display method: retaining the contextual client-side script that these UI components require for a normal execution, and only setting the irrelevant content from this client-side script as hidden.

Another important consideration in large-scale reuse is the security issue. As the proposed approach extracts content from general Web applications, the content from malicious webpages could be included. However, as the proposed approach does not directly reuse the runnable JavaScript code, the content-based reuse is not as vulnerable as script-based reuse. For UI components extracted from potentially malicious webpages, the proposed approach actually provides a protection mechanism by dividing the searched content or UI components into isolated subpages [28] called iframes. For different iframes, different permissions are granted to allow them to execute only legitimate operations. Even if an iframe contains malicious content, it will not contaminate the others. The modularity of iframes in our resulting webpage mitigates the security risks.

The content searched by the proposed approach is generally taken from public sources and includes many different types, e.g., encyclopedias, articles of news, blogs and posters from discussion forum, information of items or products, government records and financial data. Among these different types, some of them are in the well-defined format (financial data like currency exchange ratio provided by an open service), while some others are not (static text and image searched from public profiles provided by a Social Networking Service (SNS)). No matter the content is in the well-defined format or not, the proposed approach mainly aims to describe and integrate the content. We are not searching for private data, usually protected by encryption or encoding, which is ignored by the system.

## V. Conclusion and Future Work

In this paper, we presented an effective approach to integrate content from Web applications and services/feeds for personal use. The proposed approach uses WCDL to describe the Web content and functionalities, and produces a lightweight integration using a Web content extraction and hide-and-display method. Using the proposed extraction and integration system, typical users can construct mashup Web applications easily and quickly. We also undertook an exploratory analysis and evaluation of UI components for large-scale reuse. This will provide reusable components in future developments.

As a future work, we plan to crawl a significant number of websites/webpages and construct a highly precise retrieval system for large-scale UI components. Moreover, we will explore additional flexible methods of integration and construct an open community for sharing mashup components.

## VI. Acknowledgement

This work was supported by a Grant-in-Aid for Young Scientists B (No.26730158) from the Japan Society for the Promotion of Science (JSPS).

## References

[1] Yahoo Pipes, URL: http://pipes.yahoo.com/pipes/ [accessed: 2015-02-15].

[2] J. Yu, B. Benatallah, R. Saint-Paul, F. Casati, F. Daniel, and M. Matera, “A framework for rapid integration of presentation components,” in The Proceedings of the 16th International Conference on World Wide Web, 2007, pp. 923–932.

[3] J. Yu, B. Benatallah, F. Casati, and F. Daniel, “Understanding mashup development and its differences with traditional integration,” Internet Computing, vol. 12, no. 5, 2008, pp. 44–52.

[4] J. Tatemura, A. Sawires, O. Po, S. Chen, K. S. Candan, D. Agrawal, and M. Goveas, “Mashup feeds: Continuous queries over Web services,” in The Proceedings of the 33th SIGMOD International Conference on Management of Data, 2007, pp. 1128–1130.

[5] M. Sabbouh, J. Higginson, S. Semy, and D. Gagne, “Web mashup scripting language,” in The Proceedings of the 16th International Conference on World Wide Web, 2007, pp. 1035–1036.

[6] E. M. Maximilien, H. Wilkinson, N. Desai, and S. Tai, “A domain-specific language for Web apis and services mashups,” in The Proceedings of the 5th international conference on Service-Oriented Computing, 2007, pp. 13–26.

[7] E. Wohlstadter, P. Li, and B. Cannon, “Web service mashup middleware with partitioning of XML pipelines,” in The Proceedings of 7th International Conference on Web Services, 2009, pp. 91–98.

[8] Q. Zhao, G. Huang, J. Huang, X. Liu, and H. Mei, “A Web-based mashup environment for on-the-fly service composition,” in The Proceedings of 4th International Symposium on Service-Oriented System Engineering, 2008, pp. 32–37.

[9] J. Fujima, A. Lunzer, K. Hornbaek, and Y. Tanaka, “C3W: clipping, connecting and cloning for the Web,” in The Proceedings of the 13th International World Wide Web conference, 2004, pp. 444–445.

[10] J. Wong and J. I. Hong, “Making mashups with marmite: Towards end-user programming for the Web,” in The Proceedings of the SIGCHI Conference on Human factors in computing systems, 2007, pp. 1435–1444.

[11] XML User Interface Language, URL: http://developer.mozilla.org/en-US/docs/Mozilla/Tech/XUL [accessed: 2015-02-15].

[12] R. Ennals and M. Garofalakis, “MashMaker: Mashups for the masses,” in The Proceedings of the 33th SIGMOD International Conference on Management of Data, 2007, pp. 1116–1118.

[13] Creating Web Clip Widgets, URL: https://support.apple.com/kb/PH14092 [accessed: 2015-02-15].

[14] F. Daniel and M. Matera, “Turning Web applications into mashup components: Issues, models, and solutions,” in The Proceedings of the 9th International Conference on Web Engineering, 2009, pp. 45–60.

[15] J. Guo and H. Han, “Analysis and design of programmatic interfaces for integration of diverse Web contents,” International Journal of Software Engineering and Knowledge Engineering, vol. 23, no. 10, 2013, pp. 1487–1511.

[16] Java Portlet, URL: http://www.jcp.org/en/jsr/detail?id=286 [accessed: 2015-02-15].

[17] Web Services for Remote Portlets, URL: http://www.oasis-open.org/committees/wsrp/ [accessed: 2015-02-15].

[18] Simple Object Access Protocol, URL: http://www.w3.org/TR/soap/ [accessed: 2015-02-15].

[19] Representational State Transfer, URL: http://rest.elkstein.org/ [accessed: 2015-02-15].

[20] Yahoo Query Language, URL: http://developer.yahoo.com/yql/ [accessed: 2015-02-15].

[21] Web Service Definition Language, URL: http://www.w3.org/TR/wsdl [accessed: 2015-02-15].

[22] HtmlUnit, URL: http://htmlunit.sourceforge.net/ [accessed: 2015-02-15].

[23] H. Han and T. Tokuda, “A method for integration of Web applications based on information extraction,” in The Proceedings of the 8th International Conference on Web Engineering, 2008, pp. 189–195.

[24] HTML DOM Script Object, URL: http://www.w3schools.com/jsref/dom_obj_script.asp [accessed: 2015-02-15].

 ISBN: 978-1-61208-392-6 

[25] Web Application Description Language, URL: https://wadl.dev.java.net/ [accessed: 2015-02-15].

[26] News Search Documentation for Yahoo! Search, URL: http://developer.yahoo.com/search/news/V1/newsSearch.html [accessed: 2015-02-15].

[27] H. Han, P. Gao, Y. Xue, C. Tao, and K. Oyama, "Analysis and design: Towards large-scale reuse and integration of Web user interface components," in Information Reuse and Integration In Academia And Industry. Springer Verlag, 2013, pp. 133–162.

[28] Y. Cao, V. Yegneswaran, P. Porras, and Y. Chen, "A path-cutting approach to blocking XSS worms in social Web networks," in The Proceedings of the 18th ACM conference on Computer and communications security, 2012, pp. 745–748.

 ISBN: 978-1-61208-392-6 

# Enhancing Object-Oriented UML for Developing an Intelligent Semantic Knowledge Dictionary

*Mansoor Al Aali, Samia Yousif*
*College of Information Technology*
*Ahlia University*
*Manama, Kingdom of Bahrain*
*Emails: { malaali, syousif }@ahlia.edu.bh*

***Abstract*—This research presents a new SEMANTIC dictionary system developed by utilizing the Object-Oriented (OO) approach. SEMANTIC dictionary system contains English words, their Arabic meanings, associated actions, semantic relationships, inherited actions and attributes, exceptional relationships and other semantics and characteristics. SEMANTIC dictionary system utilizes Object-Oriented major features, such as objects, classes, aggregation, inheritances, encapsulation and polymorphism. SEMANTIC dictionary is a knowledge base, which can be considered as an intelligent language model and can be used for many language teaching purposes. This research shows how simple phrases can be generated or validated to be semantically correct. In the process of using OO UML to represent semantic knowledge, we have made enhancements and additions to UML itself.**

***Keywords-Semantic; Dictionary; English; Object Orientation; Teaching***

## I. INTRODUCTION

Research into natural language processing has received great attention and will continue to do so because of the importance of such topic to humanity [1][18]. Natural language processing presents two main challenges: representing semantics and querying these semantics. Until now, no one claims to have the solution to store semantics of complex sentences and link all its words together with the meaning of words, phrases of other sentences [2][12]. This research presents a novel approach by storing the semantics of English simple phrases using the Object Oriented techniques for teaching and learning English. Object oriented technique is a method of viewing knowledge at different levels of abstraction [19].

English is the first or second language in most countries educational systems. Students in many countries study the English language starting from the 1st grade of elementary schools [9][13][14]. Mastering English in both writing and reading is essential and crucial to the success of students in high schools and colleges for them to be able to find a good job in the future [11]. Furthermore, the students need to feel more of the knowledge relating to the words to improve their writing skills in English. Students studying English as Second Language (ESL) need to use dictionaries in order to know what words mean and what they can do with them in order to understand and communicate in English [6].

Dictionaries are one of many important tools [10] that can increase the students' understanding of the language and develop their vocabulary. Furthermore, Learners' Dictionaries are used to practice English correctly, effectively and appropriately [17]. Consequently, dictionaries must have maximum semantic information for words related to semantic classification that will assist the students to understand English language, as well as improve and enhance their writing in English.

Several developers have attempted to develop inclusive and unique dictionaries with special features that may help the dictionary to become more useful and popular. We have conducted a survey on online dictionaries and elicited that almost all dictionaries come under three categories: leaner, translator and search. Each dictionary has a number of features, for example, some dictionaries provide words definition, examples, pronunciation, words translation, text translation, pictures, rhymes, synonyms and words functions like noun, adjective, etc. Many companies have cashed in on the concept and have developed dictionaries for a variety of purposes such as Cambridge Learner's Dictionary, Word Central, Ultralingua Language Software, RhymeZone, Wordsmyth Children's Dictionary, McGraw-Hill Children's Dictionary, OneLook Dictionary Search, Freesearch, Sakhr, Tarjim, WordNet and ALWafi.

Writing fluently in English language requires knowledge of the conventional contexts, strong vocabulary, and collocations surrounding the word. While this information may be presented implicitly in dictionaries to help ESL learners, ESL students often have difficulties finding the right meaning for unfamiliar words or phrases in their dictionaries. In response to these difficulties, some researchers have proposed new ways to access the English lexicon [8] and presenting typical phraseology rather than words in isolation [6]. In addition, some researchers have suggested a representation, with a tree-based model of runtime dictionaries, allowing a very fast morphological analysis with semantic recognition, for example, the path <SPORTS/Water Sports/Swimming> describes the document field <Swimming> as a sub-field of <Water Sports>, which is a sub-field of super-field <SPORTS> [3].

This research presents some aspects of semantics for simple sentences and phrases, for example, the simple phrase (boy eats book) is grammatically correct while it is wrong in real world meaning. In turn, SEMANTIC dictionary system

 ISBN: 978-1-61208-392-6 

rejects this sentence and displays the reasons for the rejection and provides a dynamic correct example such as (boy eats apple). This is made possible because of the intelligent features of the Object-Oriented approach and the actions associated with each word (noun). Some researchers proposed a practical retrieval system for semantic representation by building relationships between verbs and nouns in derivation frames [3]. In SEMANTIC dictionary, administrators create relationships between actions (verbs) and classes (nouns) to provide important information about the meanings of nouns. For instance, the action 'eat' could be the relationship between the super classes 'human' and 'food' which means 'human eat food'.

Teaching English needs to be changed or improved by adding various tools and new strategies that will assist students to comprehend English in the best way [5]. Many researchers continue to study a variety of new approaches in teaching English [1][2][7][11][12][13][14][15]. Communicative language teaching approach advocates the development of communicative competence as a primary goal through the extensive use of the second language as a means of communication during classroom lessons. Using computers raises students' motivation and their performance as well as placing students in a learning state with full participation, concentration and enjoyment [4][5].

SEMANTIC dictionary system is a web-based dictionary system. Each word in this dictionary belongs to a class and may have one or more subclasses. Subclasses inherit all the public attributes and operations of their super class and this concept is utilized in all types of processing on the SEMANTIC dictionary system. For example, the super class 'human' inherits the public operation 'eat' to its subclass 'boy' and therefore we can generate or check the simple sentence 'boy eats food' or 'boys eat food' with semantic meaning.

The work in this research signifies that SEMANTIC dictionary system assists teachers to make the English subject more interesting and raise student's motivation and contribution, as well as it enables students to understand the meaning of words and helps them to organize sentences and build them correctly with semantic meaning.

The rest of the paper consists of a number of sections. Section II presents the technique for incorporating semantics into a dictionary using the object oriented approach. In section II we present a summary for enhancing UML code and UML diagram incorporating semantics in the SEMANTIC Dictionary System. The rest of the paper presents the different types of associations between objects using the enhanced UML code such as Normal Association (Public Operations), Recursive Association (Public Operations), Private Association (Private Operations), Exceptional Association (Exceptional Operations) and Subject Only Operations. We end the paper with our conclusions.

Representing language semantics has been considered using a variety of tools, techniques and theories, eg, using the Web Ontology Language (OWL) [20][21]. OWL has been accepted by the World Wide Web Consortium (W3C) as the standard knowledge representation formalism for the Semantic Web [21]. OWL provides a very expressive selection of syntactic constructs for Semantic Web ontology development. There are many other techniques used over time but it is outside the scope of this paper to survey them. The main focus of the research presented in this paper is to represent simple language semantics extracted from simple phrases or sentences to be used as a dictionary of knowledge mainly for teaching English and for English to Arabic translation. A number of researchers have studied representing Arabic semantics [22]. We have chosen UML [24] for its simplicity and it is suitable for the representation requirements. Our research is not concerned with web semantics nor with general knowledge representation, but concerned primarily with representing simple Arabic and English phrases for teaching English simple sentences and for translation and we demonstrate its suitability in this paper. Section II presents the enhancements to UML necessary to incorporate the required level of semantics for the SEMANTIC dictionary system.

The rest of the paper is divided into a number of sections. Section II presents the approach of applying the object-oriented techniques to incorporate semantics in the semantic dictionary. Section II presents the Enhancements to UML for SEMANTIC Dictionary System and the enhancement of UML diagram for SEMANTIC dictionary system. Section III presents a summary discussion for the subject-only concept and supporting scenarios. Section IV presents the conclusions and future work.

## II. APPLYING THE OBJECT-ORIENTED TECHNIQUES TO INCORPORATE SEMANTICS IN THE SEMANTIC DICTIONARY

Object-Oriented (OO) is a paradigm for creating software systems using objects. Objects are tangible and conceptual things we find in the real world. Using OO makes SEMANTIC dictionary system objects more semantically related and hence better intelligence can be implemented. OO has major notions such as objects, classes, inheritances, encapsulation and polymorphism. The OO Unified Modeling Language (UML) is the industry-standard language for specifying, visualizing, constructing, and documenting the artifacts of software systems. It simplifies the complex process of software design, making a "blueprint" for construction. These OO concepts were used as a model for representing SEMANTIC knowledge base rather than SEMANTIC development. The main reason for using UML is to model the simple semantics of the language phrases which can then be modelled internally in our software system depicting the SEMANTIC DICTIONARY system. Modelling using UML means the ability to directly represent the knowledge using an Object oriented language. The reasons for expanding or extending the UML is that the original UML does not have certain semantic concepts suitable for our purpose as will be illustrated in the next section.

 ISBN: 978-1-61208-392-6 

### A. Enhancing UML for SEMANTIC Dictionary System

The public attributes and operations are expressed with a plus sign (+). The private attributes and operations are expressed with a minus sign (-). The exceptional operations are expressed with a less-than sign (<). The "Subject-Only" operations are expressed with (&). The letter "b" is added to these signs (+, -, < or &) when the type of the operation is "Action By" while the letter "o" is added to them when the type of the operation is "Action On". Notice that both "Exceptional" and "Subject-Only" operations are not a part of UML class diagram, but we added them, which means that the signs are according to our view as well as the addition of the letters "b" and "o" to the signs. The signs and the letters are utilized to distinguish between the actions.

### B. Enhancement of UML Diagram for SEMANTIC Dictionary System

Figure 1 demonstrates two super classes with their subclasses, instances and relationships. The dotted lines mean that the diagram can be extended depending on the system. The first super class is called Super class A and it has Subclass Aa. This subclass contains three instances: Instance Aa1, Instance Aa2 and Instance Aa3. The second super class is called Super class B and it has Subclass Ba. This subclass consists of three instances: Instance Ba1, Instance Ba2 and Instance Ba3. The super classes inherit the public operations (Actions) to its subclasses. Accordingly, these subclasses inherit its public Actions to its subclasses or instances. Super classes are illustrated with rectangles divided into two compartments. Place the name of the class in the first partition (centered, bolded, and capitalized each word), list the subclasses or instances in the second partition. The following are the relationships between classes and the drawings of these relationships, as shown in Figure 2.

**Normal Association (Public Operations):** It is a connection between classes, a semantic connection between objects (instances) of the classes involved in the association. It is normally bidirectional, which means that if an object is associated with another object, both objects are aware of each other. First direction is Action By (+b Action), and second direction is Action On (+o Action). The multiplicity (0 .. *) and (1 .. *) are used to present how many objects are linked. This association is inherited to the subclasses. It is sketched by a wide straight line connected between two super classes. The Action By (+b Action) is put above this line while the Action On (+o Action) is placed below it. Moreover, the multiplicity and directions' arrows are presented on this line.

**Recursive Association (Public Operations):** It is a connection between a class and itself. The association still presents a semantic connection between objects, but the connected objects are of the same class. This is also bidirectional and it has multiplicity like normal association. This association is inherited to the subclasses. It is expressed by a wide elbow connecter between the super class and itself. The Action By (+b Action) is placed above this connecter whereas the Action On (+o Action) is put below it. Furthermore, the multiplicity and directions' arrows are presented on this connecter.

**Private Association (Private Operations):** It is a connection between objects (instances). It is bidirectional; First direction is Action By (-b Action), and second direction is Action On (-o Action). This association can't be inherited. It is sketched by a thin elbow connecter between two instances. The Action By (-b Action) is presented above this connecter while the Action On (-o Action) is placed below it. In addition, directions' arrows are demonstrated on this connecter.

**Exceptional Association (Exceptional Operations):** It is a connection between some objects and classes. Exceptional Action By (<b Action) means that the inherited action that the object can't do and Exceptional Action On (<o Action) means that the inherited action can't be carried out on the object. Consequently these actions are excluded from this object. It is demonstrated by a thin straight line between the instance of a super class and another super class, the short vertical line is put before the end up of this straight line. The Action By (<b Action) is shown above the straight line whereas the Action On (<o Action) is put below the straight line.

**Subject-Only Operations:** It means that the action is done by the subject (class) without an object. It is always Action By (&b Action) as well as it is inherited to the subclasses. It is sketched by a wide straight line connected to the super class and end up by oval arrow. The Action By (&b Action) is placed above this line.

In SEMANTIC dictionary system, the relationships or associations (public operation Action By (+b Action)) are created between two super classes manually; subsequently the operations Action On (+o Action) are generated between them automatically. The normal line is sketched between two super classes (Super Class A and Super Class B) indicates the public relationship (Action By) which is created manually between them while the dotted lines are sketched between them and its classes signify the public relationship (Action On) which is generated automatically as well as it shows the inherence public operations Action By (+b Action) and Action On (+o Action) as illustrated in Figure 2.

For example, Super class A inherits the public operations (Action By and Action On) to its Subclass Aa and its instances (Instance Aa1, Instance Aa2 and Instance Aa3). Notice that the sketched lines of relationships between classes in Figure 2 are hidden in SEMANTIC dictionary system. The other relationships that were applied in SEMANTIC dictionary system can be exhibited in the same manner.

Mainly, English simple sentence consists of three major parts are The Subject, The Verb and The Object. Now, will describe the Public Associations (Public Operations) and will explain how these relationships are achieved in the SEMANTIC Dictionary System as well as will present the intelligent aspects of it. Furthermore, it will provide more details about enhanced UML for drawing this type of relationships. The administrator or the teacher will create the public relationships (associations or operations) between the super classes utilizing simple present verbs manually. SEMANTIC Dictionary System has some intelligent aspects, for example, super classes will inherit these public

 ISBN: 978-1-61208-392-6 

relationships within present verbs to its subclasses and it will produce other cases of public relationships via different verbs, such as: past verbs, future verbs and present continuous verbs automatically. The super classes will inherit these produced cases of public relationships to its subclasses as well.

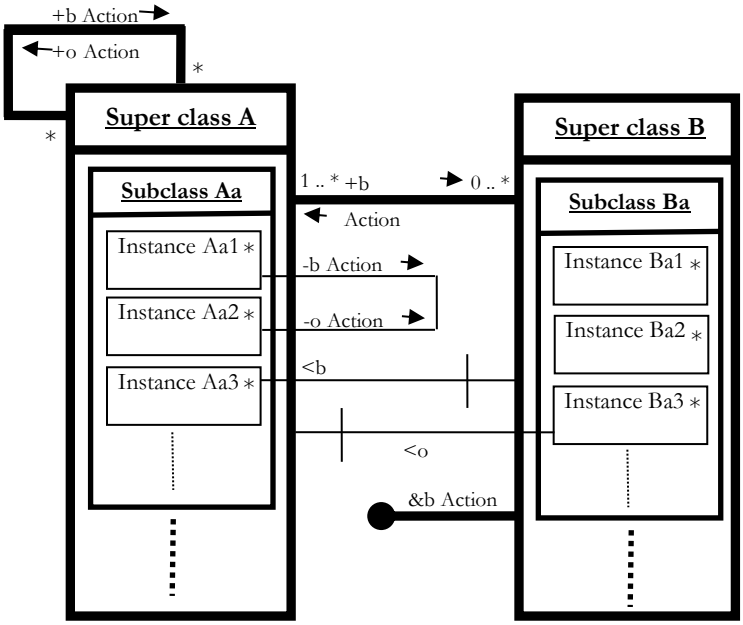



Figure 1. Enhancement UML Diagram (Meta Diagram)

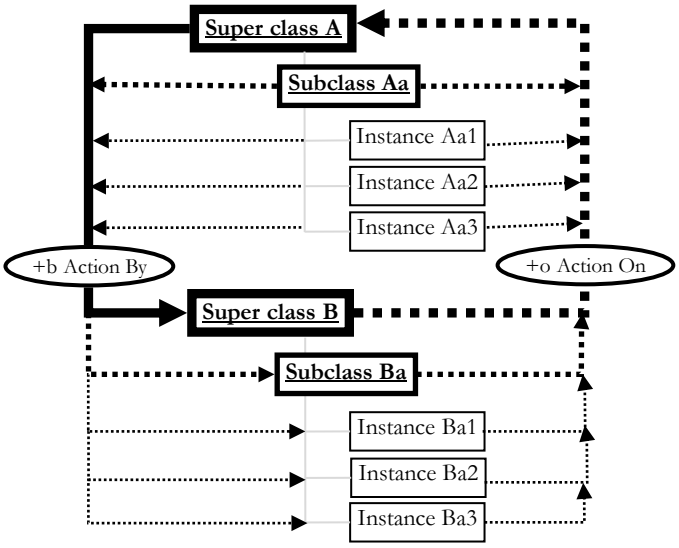



Figure 2. Classes Relationships (Meta Diagram)

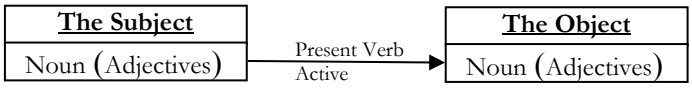



Figure 3. English Simple Sentence Meta Diagram (Active Simple Sentence within Present Verb)

 ISBN: 978-1-61208-392-6 

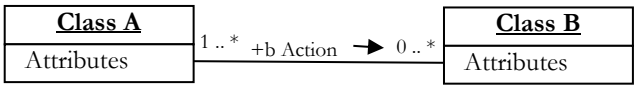



Figure 4. Enhanced UML Meta Diagram within Action By (Public Relationship via Present Verb)

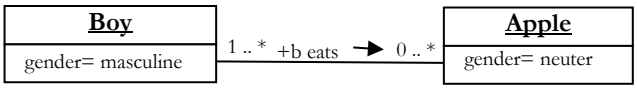



Figure 5. Example of Enhanced UML Diagram within Action By (Public Relationship via Present Verb) in SEMANTIC Dictionary System

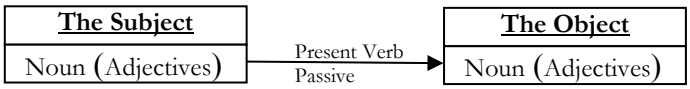



Figure 6. English Simple Sentence Meta Diagram (Passive Simple Sentence within Present Verb)

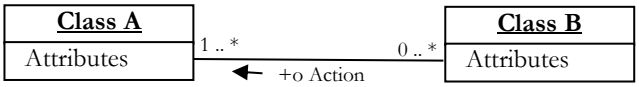



Figure 7. Enhanced UML Meta Diagram within Action On (Public Relationship via Present Verb, which is produced)

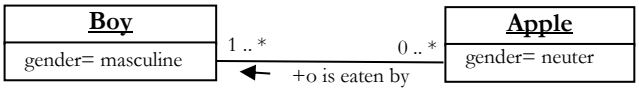



Figure 8. Example of Enhanced UML Diagram within Action On (Public Relationship via Present Verb) in SEMANTIC Dictionary System

Consequently, SEMANTIC Dictionary System will generate several English simple sentences automatically, for example: active simple sentences and passive simple sentences within present verbs, active simple sentences and passive simple sentences within past verbs, active simple sentences and passive simple sentences within future verbs, active simple sentences and passive simple sentences within present perfect verbs and active simple sentences within present continues verbs. In addition, SEMANTIC dictionary system will generate these sentences using singular subject, plural subject, singular object and plural object. The following demonstrates the public relationships utilizing present verbs to generate active simple present sentence.

Figure 3 illustrates the Meta diagram of active simple sentence within a present verb. Depending on Figure 3, we have drawn the enhanced UML Meta diagram; see Figure 4. One of "Class A" or more can carry out the Action By (+b Action) on one of "Class B" or more. On other hand, one of "Class A" or more can perform the Action By (+b Action). Here, the Action By is in the present. Similarly, "Class A" and "Class B" have many attributes. Figure 5 is an example in the SEMANTIC Dictionary System; a Boy is a human and his gender is masculine, an Apple is a vegetarian food and its gender is neuter. The public relationship between the Boy and the Apple is the Action By, which is the present verb (eat), that they have already inherited from their super classes (Human and Vegetarian Food respectively). This is the intelligent aspect of the SEMANTIC dictionary system. Hence, SEMANTIC will generate numerous sentences automatically, as presented in the following: A boy eats, Boys eat, A boy eats an apple, A boy eats apples, Boys eat an apple, Boys eat apples.

Passive simple present sentence means that "The Subject" does not perform the action of the present verb in the sentence - the action is performed on it. Figure 6 illustrates the Meta diagram of passive simple sentence within a present verb. Regularly, "The Subject" and "The Object" are nouns and they have adjectives, whereas when we write a simple sentence, we can write "The Subject" and "The Object" without any adjective. Moreover, we can write them in singular and plural.

Depending on Figure 6, we have drawn the enhanced UML Meta diagram, see Figure 7. Action On (+o Action) is the action that can be carried out on one of "Class B" or more by one of "Class A" or more. Similarly, "Class A" and "Class B" have many attributes. Figure 8 is an example in the SEMANTIC Dictionary System; a Boy is a human and his gender is masculine, an Apple is a vegetarian food and its gender is neuter. The super classes (Human and Vegetarian Food) inherit the public relationship

 ISBN: 978-1-61208-392-6 

(present verb: eat) to their subclasses (the Boy and the Apple respectively). Since SEMANTIC dictionary system has intelligent aspects, it will produce the Action On between the subclasses (the Boy and the Apple), which is the action (eaten by). Furthermore, SEMANTIC will generate various sentences automatically, as presented in the following: An apple is eaten, Apples are eaten, An apple is eaten by a boy, An apple is eaten by boys, Apples are eaten by a boy, Apples are eaten by boys.

## IIII. SUBJECT-ONLY EXPLANATION AND SCENARIOS

The "Subject-Only" operations are expressed with an ampersand (&). The letter "b" is added to these signs (+, -, < or &) when the type of the operation is "Action By" while the letter "o" is added to them when the type of the operation is "Action On". Notice that both "Exceptional" and "Subject-Only" operations are not a part of the original UML class diagram, but we added them, which means that their signs are according to our view, as well as the addition of the letters "b" and "o" to the signs. The signs and the letters are utilized to distinguish between the actions. In addition, Figure 9 presents the following:

- **Classes:** SEMANTIC dictionary system has several super classes and many subclasses whereas Figure 9 presents three Super classes: Human, Vegetarian Food and Information Media. In addition, it shows six subclasses, such as (Reading) of super class Information Media and (Family and People) of super class Human.
- **Instances:** In SEMANTIC dictionary system, each subclass has many instances, for example (Man, Woman, Boy and Girl) of subclass People. Another example: the subclass Fruit has many instances that are applied in SEMANTIC dictionary system while Figure 9 shows (Apple and Banana) as an example.
- **Aggregation:** A lot of aggregations are applied in SEMANTIC dictionary system, for example the super class Information Media consists of class Reading and the class Reading consists of objects or instances (Book, Magazine and Newspaper).
- **Public Operations:** SEMANTIC dictionary system has many public operations (Action By): for example; buys, sells and browses, for example: a girl buys an apple. As well as the public operations (Action On) which are generated automatically: such as; bought by, sold by and browsed by, for example: an apple is bought by a girl.
- **Private Operations:** SEMANTIC dictionary system has a private operation (Action By): "engages" for object Man. For example: a man engages a woman and a private operation (Action On): "engaged to" for object Woman. For example: a woman engaged to a man.
- **Subject_Only Operations:** SEMANTIC dictionary system contains several Subject_Only operations (Action By): such as Sits, Grows, Sleeps, Swims and Weeps. For example, a boy sits.
- **Exceptional operations:** SEMANTIC dictionary system has an exceptional operation (Action By): Cooks. For example, a boy cannot cook a cloth. Also, an exceptional operation (Action On): Cooked by. For example, a football is not cooked by a human.
- **Inheritance:** In SEMANTIC dictionary system the super classes inherit the public operations (Actions) to its subclasses. For example: the super class Human inherits its public operations such as (eat, cook, read and write) to its subclasses (People and Family) and consequently these classes inherits these operations to its subclasses or instances like (Man, Woman, and Boy).
- **Multiplicity:** Different multiplicities are applied in SEMANTIC dictionary system. For example: "a man eats an apple", "Three men eat an apple", "a man eats two apples" and "four men eat five apples".
- **Normal Association:** SEMANTIC dictionary system has many normal associations, for example, first direction: a boy reads many (zero or many) books. Second direction: a book can be read by many (one or many) boys.
- **Recursive Association:** SEMANTIC dictionary system has a recursive association, first direction: (one or many) woman feeds (one or many) sons. Second direction: (one or many) son fed by (one or many) women.

## IV. CONCLUSION AND FUTURE WORK

This paper presented a new approach for the development and evaluation of an intelligent English-Arabic Object Oriented dictionary system called SEMANTIC. This research is the first to effectively use the OO concept and enhance UML to deal with English language semantics for simple phrases. SEMANTIC Dictionary System was tested on the 6$^{th}$ class of an elementary school for girls in kingdom of Bahrain to test its effectiveness in teaching some aspects of the English language and to test its level of contribution to students learning and the tests are very promising [23].

 ISBN: 978-1-61208-392-6 

We are planning to compare our enhanced UML with OWL for the representation of natural language semantics especially for the application of translation from English to Arabic.

ACKNOWLEDGEMENT

We would like to thank Ahlia University for providing the support and our students for working on developing and testing some of our research ideas.

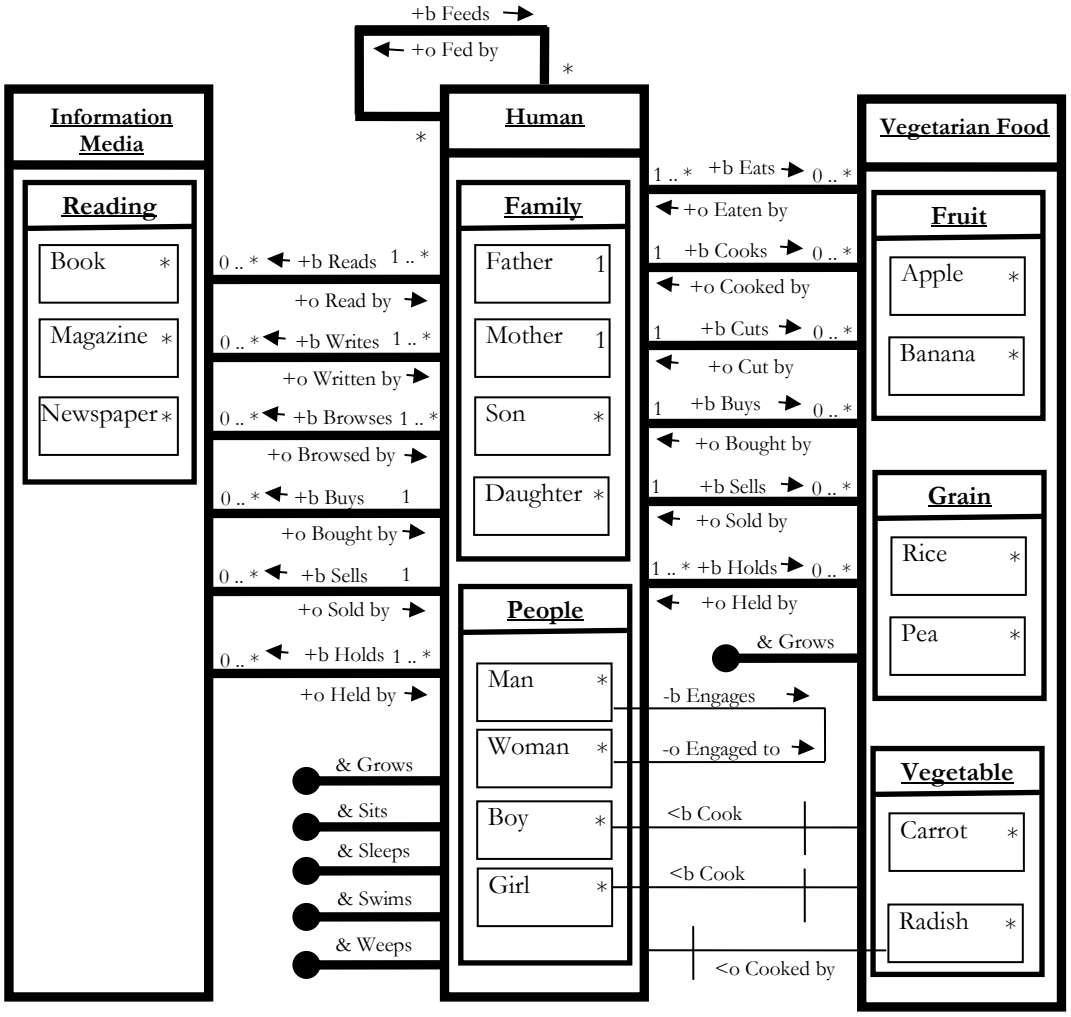



Figure 9. Example of Enhanced UML Diagram in the SEMANTIC Dictionary System.

 ISBN: 978-1-61208-392-6 

## REFERENCES


[1] Al-A'ali Mansoor, "An Approach for Automatically Extracting Translation Rules For a Better English-to-Arabic Machine Translation", Journal of Computer Science, Science Publishers, USA, 2007, pp. 1228-139.

[2] Al-A'ali Mansoor, "An Efficient Technique for Identifying and Storing Syntactic and Semantic English Natural Language Sentence Patterns", WSEAS Transactions On Systems, vol. 6, Issue 1, 2007, pp. 173-178.

[3] El-Sayed Atlam, K Morita, M Fuketa, and Jun-ichi Aoe, "A new method for selecting English field association terms of compound words and its knowledge representation", Information Processing & Management, vol. 38, Issue 6, 2002, pp. 807-821.

[4] Gwo-Jen Hwang, Po-Han Wu, and Chi-Chang Chen, "An online game approach for improving stu-dents' learning performance in web-based problem-solving activities", Computers & Education, vol. 59, 2012, pp. 1246–1256.

[5] Huiwei Cai, "E-learning and English Teaching. IERI Procedia", vol. 2, 2012, pp. 841-846.

[6] Janice McAlpine and Johanne Myles, "Capturing phraseology in an online dictionary for advanced users of English as a second language: a response to user needs", System, vol. 31, Issue 1, 2003, pp. 71-84.

[7] Mansoor Al-Aali and Samia Yousif, "An Intelligent Multi-Lingual Object Oriented Dictionary System", Proceedings of the 11th WSEAS International Conference on SYSTEMS, Agios Nikolaos, Crete Island, Greece, July 23-25, 2007, pp. 468-473.

[8] Mark A. Changizi, "Economically organized hierarchies in WordNet and the Oxford English Dictionary", Cognitive Systems Research, vol. 9, 2008, pp. 214–228.

[9] Masoud Hashemi and Masoud Azizinezhad, "Teaching English To Children: A Unique, Challenging Experience For Teachers, Effective Teaching Ideas", Procedia - Social and Behavioral Sciences, vol. 30, 2011, pp. 2083-2087.

[10] Octavio Santana Suárez, Francisco J. Carreras Riudavets, Zenón Hernández Figueroa, and Antonio C. González Cabrera, "Integration of an XML electronic dictionary with linguistic tools for natural language processing", Information Processing and Management, vol. 43, 2007, pp. 946-957.

[11] Sally Wu, "Knowledge based approaches for English-Chinese interpretation teaching and learning", Procedia Social and Behavioral Sciences, vol. 2, 2010, pp. 2383-2387.

[12] Samia Yousif and Mansoor Al-A'ali, "Object Oriented Techniques for an Intelligent Multi-Purpose English Language Dictionary System", 9th ICEIS – NLPCS 2007, Funchal-Madeira/Portugal, 2007, pp. 246-255.

[13] Senem Seda Sahenk Erkan, "Problems of English language instruction in primary in Turkey and their suggestions", Procedia - Social and Behavioral Sciences, vol. 46, 2012, pp. 1117-1121.

[14] Serife Demircioglu, "Teaching English vocabulary to young learners via drama", Procedia Social and Behavioral Sciences, vol. 2, 2010, pp. 439-443.

[15] Shehdeh Fareh, "Challenges of teaching English in the Arab world: Why can't EFL programs deliver as expected", Procedia Social and Behavioral Sciences, vol. 2, 2010, pp. 3600–3604.

[16] Shireen Motala, "Quality and indicators of quality in South African education: a critical appraisal", International Journal of Educational Development, vol. 21, 2001, pp. 61-78.

[17] Stefania Nuccorini, "Pragmatics in learners' dictionaries", Journal of Pragmatics, vol. 19, Issue 3, 1993, pp. 215-237.

[18] Taher S.K. Homeed and Mansoor Al-A'ali, "Analyzing Statistical and Syntactical English Text For Word prediction and text generation", journal of Information Technology, ANSI, vol. 6, issue 7, 2007, pp. 954-962.

[19] Xueyun Ji, "Analysis and Design for Object-oriented Multi-tier Architecture of Public Opinion Survey System Based on UML", Procedia Engineering, vol. 15, 2011, pp. 5445-5449.

[20] Mohd Syazwan Abdullah, Richard Paige, Ian Benest, and Chris Kimble, "Knowledge Modelling Using The UML Profile.", in Ilias Maglogiannis; Kostas Karpouzis & Max Bramer, ed., 'AIAI' , Springer, 2006, pp. 70-77.

[21] Cartik R. Kothari, "Enhancing Owl Ontologies With Relation Semantics", International Journal Of Software Engineering and Knowledge Engineering, vol. 18, no.3, 2008, pp. 327–356.

[22] Bassam Haddad, "Semantic Representation of Arabic a Logical Approach Towards Towards Compositionality and Generalisedx Arabic Quantifiers", 2007, International Journal of Computer Processing of Oriental Languages, vol. 20, no. 1, pp. 38-52.

[23] Samia Yousif and Mansoor Al Aali, Evaluation of SEMANTIC: An Intelligent English Semantic Object Orientated Dictionary System, CONTENT 2015, The Seventh International Conference on Creative Content Technologies, March 22 - 27, 2015. Nice, France, to appear.

[24] Hassan Gomaa, Designing Concurrent, Distributed, and Real-Time Applications with UML, ISBN: 9780321951816, Jan 2014.


 ISBN: 978-1-61208-392-6 

# *Evaluation of SEMANTIC:*
## An Intelligent English Semantic Object-Oriented Dictionary System

*Samia Yousif, Mansoor Al Aali*
*College of Information Technology*
*Ahlia University, Manama, Kingdom of Bahrain*
*Emails: {syousif, malaali}@ahlia.edu.bh*

***Abstract*—There has been a great deal of research on natural language representation and processing, but little or no significant research on building intelligent semantic dictionary systems. This paper presents research on evaluating our intelligent English semantic dictionary system and its suitability for teaching English language to school children and for its use as a translation system, especially for non-native English speakers. SEMANTIC dictionary system is an intelligent Object-Oriented (OO) eLearning system based on advanced UML. It contains English words, their translation to Arabic, associated actions, semantic relationships, inherited actions and attributes, special and exceptional relationships and other semantics and characteristics. SEMANTIC dictionary system was tested on the 6th class of an elementary school for girls in the Kingdom of Bahrain to evaluate its effectiveness in teaching English and for its level of contribution to students learning. The analysis of experimental results are very promising.**

***Keywords-eLearning; Semantic; English; Object-Oriented, Evaluation.***

## I. INTRODUCTION

eLearning means both teachers and students utilize electronic media (text, audio, images, animation and video), as well as education and communication technologies [3]. In recent years, the use of new teaching methods and technology in the classroom has been increasing [6]. Teachers have used the Internet for language teaching, as well as Computer-Assisted Language Learning (CALL). The use of new technologies and teaching methods, such as Internet and CALL, enhances and complements traditional language teaching. Literature on instructional technology shows that the use of the Internet in teaching has the potential to motivate students and teachers, increase their participation, and interaction in the classroom, allow for a deeper integration with the culture of the target language, and provide students with a more active role in their learning [5][6].

Nowadays, most countries, where English is not their first language, teach English language to their students [4][9]. In turn, the traditional language teaching in the classroom needs to be integrated and enhanced by applying new teaching methods [13], like eLearning [6] and information and communication technologies. As a result, it will assist English language learners to reduce mechanical memorization and develop their English language skills [10].

Some researchers applied the ideas from CAL (Computer Assisted Language) and language learning, emphasizing context for language learning, especially for abstract word learning. Some have developed a multimedia web-based CAL system which includes 13 abstract words and five main modules: Learning Material, Testing, Communication, Help, and Extensive English Learning Web Sites [14][15]. They conducted CAL in an elementary school in southern Taiwan, and they endorse that using CAL for English teaching can enhance and improve both the quality and quantity of the English language learnt. Other researchers developed computer-assisted programs for English language teaching [11]. Their researches indicated that pupils who learnt English by applying more communicative methods made greater progress in vocabulary acquisition than those who worked with the more traditional methods [7].

Electronic dictionaries contain data in digital form. Several types of dictionaries are available electronically. There are many types of dictionaries used for different purposes. Some are used for language translation [1], while others are used for science and technology like medical and engineering dictionaries. Dictionaries are developed and continuously improved to contain a number of useful and practical applications. After thoroughly searching the literature and conducting an online dictionaries survey, we could not find any similar attempts to develop an "intelligent" dictionary to be used especially as a multi-purpose multi-lingual educational tool for young students [2][8][12].

SEMANTIC dictionary system is an intelligent web-based Object-Oriented (OO) eLearning system. Each word in it belongs to a class and may have one or more subclasses. Subclasses inherit all the public attributes and relationships (operations) of their super class and this concept is utilized in all types of processing on the SEMANTIC dictionary system. For example, consider the following words or phrase "girl read cherry" unless Cherry is the name of a book and Cherry has already read it. The simple sentence that can be composed out of these words is correct grammatically, but in a real world's semantic meaning is impossible. SEMANTIC dictionary system rejects this sentence and displays the reasons which are: Girl is human, and Cherry is food and human cannot read food. Hence, SEMANTIC dictionary system creates possible class strings which is human can read information media, and selects an object randomly from this class such as Newspaper and therefore, it corrects the semantic meaning of the simple sentence and becomes: Girl

 ISBN: 978-1-61208-392-6 

reads Newspaper. Another example: the following words are "mother feed son". SEMANTIC dictionary system accepts this simple sentence and presents the explanation which is: Mother is human, and Son is human, and human can feed human and therefore, Mother feeds Son is semantically correct. These are made possible because of the intelligent features of the OO techniques [16] and the actions (relationships) associated with each word in SEMANTIC dictionary system.

SEMANTIC dictionary system was tested on a 6th class of an elementary school for girls in the Kingdom of Bahrain to evaluate its effectiveness in teaching English, its level of contribution to students learning. The analysis of experimental results in this research shows that SEMANTIC dictionary system can help non-native young students to learn English and English semantics in a much easier manner, as well as help very young children construct simple sentences and phrases. SEMANTIC supports teachers to prepare lessons and quizzes automatically. Furthermore, it helps them to enhance the traditional methods of teaching English and thus making classrooms more interesting. SEMANTIC can also, raise students' motivation and contribution, as well as enabling students to understand the meaning of words and assists them to establish sentences correctly with semantic meaning.

Section 2 shows the Unified Modeling Language (UML) diagram for SEMANTIC ENGLISH system and how we have made enhancements and additions to UML itself. Section 3 explains how SEMANTIC dictionary helps students to learn some aspects of English language. This research presents the results of testing SEMANTIC ENGLISH system to test its effectiveness in teaching some aspects of the English language and to test its level of contribution to students' learning. The results of such testing are shown in Sections 4, 5 and 6.

## II. IMPROVING UML DIAGRAM FOR THE SEMANTIC DICTIONARY SYSTEM

The Unified Modeling Language (UML) [17] is the industry-standard language for specifying, visualizing, constructing, and documenting the artifacts of software systems. It simplifies the complex process of software design, making a 'blueprint' for construction. Figure 1 shows a Diagram of our enhanced UML Diagram to demonstrate the inheritance, aggregation and relationships of the classes in SEMANTIC dictionary system. The public attributes and operations are expressed with a plus sign (+). The private attributes and operations are expressed with a minus sign (-). The exceptional operations are expressed with a less-than sign (<). The "Subject-Only" operations are expressed with an ampersand (&). The letter "b" is added to these signs (+, -, < or &) when the type of the operation is "Action By" while the letter "o" is added to them when the type of the operation is "Action On". Notice that both "Exceptional" and "Subject-Only" operations are not a part of the original UML class diagram, but we added them, which means that their signs are according to our view, as well as the addition of the letters "b" and "o" to the signs. The signs and the letters are utilized to distinguish between the actions. In addition, Figure 1 presents the following:

- **Classes:** SEMANTIC dictionary system has several super classes and many subclasses whereas Figure 1 presents three Super classes: Human, Vegetarian Food and Information Media. In addition, it shows six subclasses, such as (Reading) of super class Information Media and (Family and People) of super class Human.
- **Instances:** In SEMANTIC dictionary system, each subclass has many instances, for example (Man, Woman, Boy and Girl) of subclass People. Another example: the subclass Fruit has many instances that are applied in SEMANTIC dictionary system while Figure 1 shows (Apple and Banana) as an example.
- **Aggregation:** A lot of aggregations are applied in SEMANTIC dictionary system, for example the super class Information Media consists of class Reading and the class Reading consists of objects or instances (Book, Magazine and Newspaper).
- **Public Operations:** SEMANTIC dictionary system has many public operations (Action By): for example; buys, sells and browses, for example: a girl buys an apple. As well as the public operations (Action On) which are generated automatically: such as; bought by, sold by and browsed by, for example: an apple is bought by a girl.
- **Private Operations:** SEMANTIC dictionary system has a private operation (Action By): "engages" for object Man. For example: a man engages a woman and a private operation (Action On): "engaged to" for object Woman. For example: a woman engaged to a man.
- **Subject_Only Operations:** SEMANTIC dictionary system contains several Subject_Only operations (Action By): such as Sits, Grows, Sleeps, Swims and Weeps. For example, a boy sits.
- **Exceptional operations:** SEMANTIC dictionary system has an exceptional operation (Action By): Cooks. For example, a boy cannot cook a cloth. Also, an exceptional operation (Action On): Cooked by. For example, a football is not cooked by a human.
- **Inheritance:** In SEMANTIC dictionary system the super classes inherit the public operations (Actions) to its subclasses. For example: the super class Human inherits its public operations such as

 ISBN: 978-1-61208-392-6 

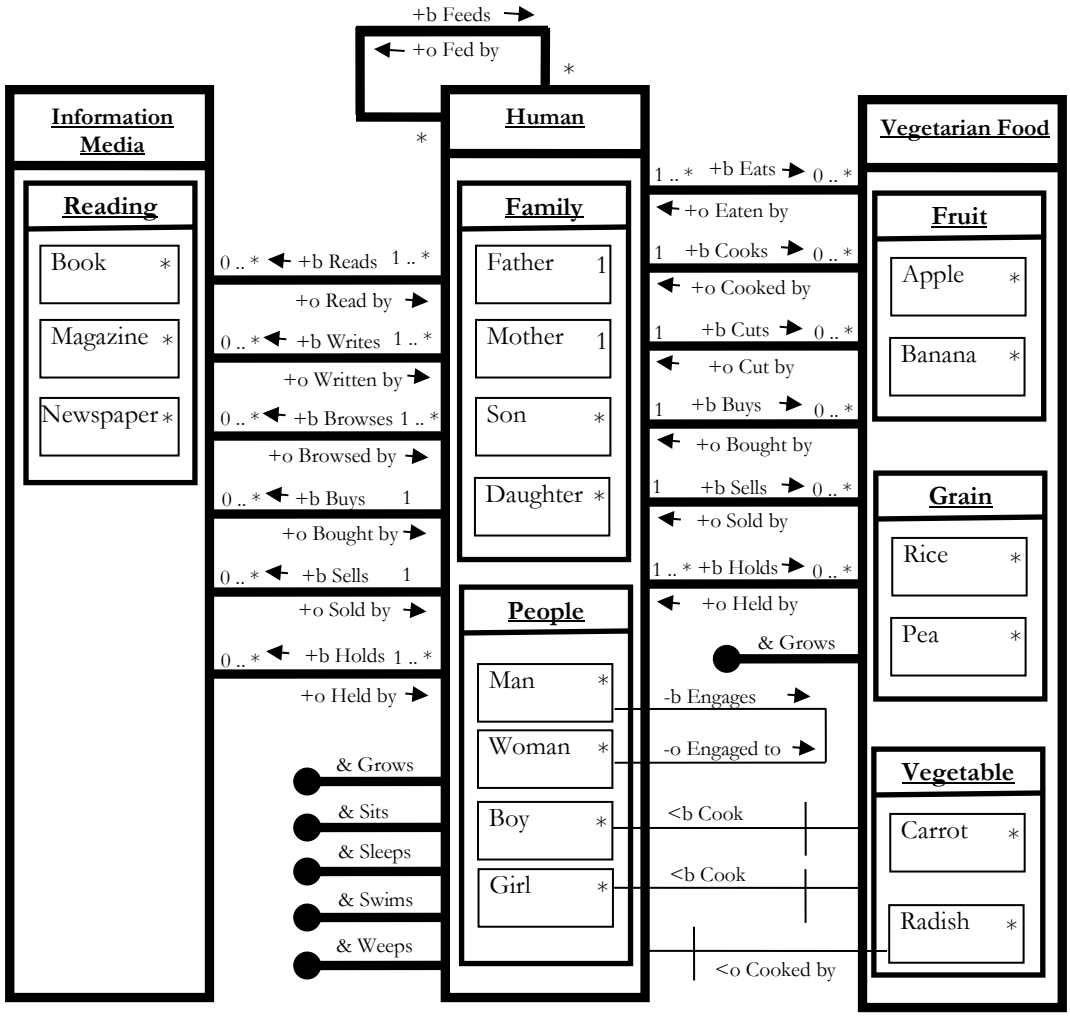



Figure 1. Example of Enhanced UML Diagram in the SEMANTIC Dictionary System.

- (eat, cook, read and write) to its subclasses (People and Family) and consequently these classes inherits these operations to its subclasses or instances like (Man, Woman, and Boy).
- **Multiplicity:** Different multiplicities are applied in SEMANTIC dictionary system. For example: "a man eats an apple", "Three men eat an apple", "a man eats two apples" and "four men eat five apples".
- **Normal Association:** SEMANTIC dictionary system has many normal associations, for example, first direction: a boy reads many (zero or many) books. Second direction: a book can be read by many (one or many) boys.
- **Recursive Association:** SEMANTIC dictionary system has a recursive association, first direction: (one or many) woman feeds (one or many) sons. Second direction: (one or many) son fed by (one or many) women.

## III. SEMANTIC FOR STUDENT LEARNING

SEMANTIC dictionary helps students to learn some aspects of English language. He can learn these aspects of English language through SEMANTIC Dictionary, lessons and Checker. SEMANTIC Dictionary contains the English words (objects) and their relationships; Action By displays a list of actions that the selected word can perform and Action On displays a list of actions which can be carried out on the selected word. The classification of the word is represented by a Tree structure. For example, when the student selects a word such as '*Boy'*, SEMANTIC Dictionary displays the classification of the word '*Boy'* which is *Thing* > *Living* > *Human* > *People* > *Boy*. This helps the student understand and recognize the semantic meaning of the word '*Boy'*. SEMANTIC dictionary also shows the associated diagram and images to provide further illustrations of the word to expand the student knowledge. Associated diagrams include information on the word in Arabic and English. The information can be plural of the word, gender, and description. Read is an action that Boy can achieve (Action

 ISBN: 978-1-61208-392-6 

By), and fed is the action that can be carried out on him (Action On). When an action is selected, SEMANTIC dictionary system displays the Arabic translation, the corresponding picture and generates simple sentence using the selected word and the action. For example, '*Boy reads*' and '*Boy is fed*'. Finally the student can listen to the pronunciation of the knowledge relating to the selected word by pressing the sound image as demonstrated in Figure 2.

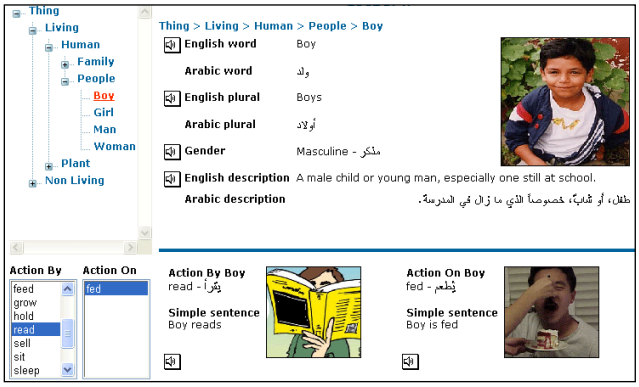



Figure 2. SEMANTIC Dictionary Tool.

SEMANTIC has three lessons; simple sentences, singular & plural nouns and the articles. SEMANTIC dictionary system explains the rules of each lesson with picture representations and examples of the words.

**Lesson1** - Simple Sentences: This lesson teaches the students how to build simple sentences with grammatical correctness. Furthermore, it teaches them how to create the simple sentences with correct semantic meaning. For example, the left side of Figure 3 presents the words utilized to build a sentence of type 2; 'the subject: Man, the verb: engage and the object: Woman', it also presents the explanation for the created sentence in the right side. SEMANTIC dictionary system accepts this sentence as indicated in Figure 3 and its explanation: '*Man is human*', and '*Woman is human*', and '*human can engage human*', and this action ('*engage'*) is special for the man; therefore '*Man engages Woman*'. SEMANTIC dictionary system follows the same manner with type 1, as shown in Figure 4.

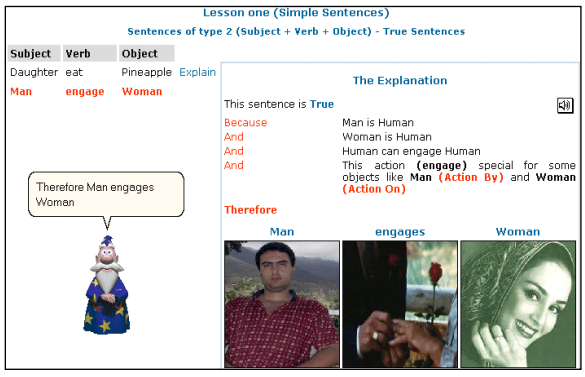



Figure 3. SEMANTIC Dictionary System accepts the created Sentence of Type 2.

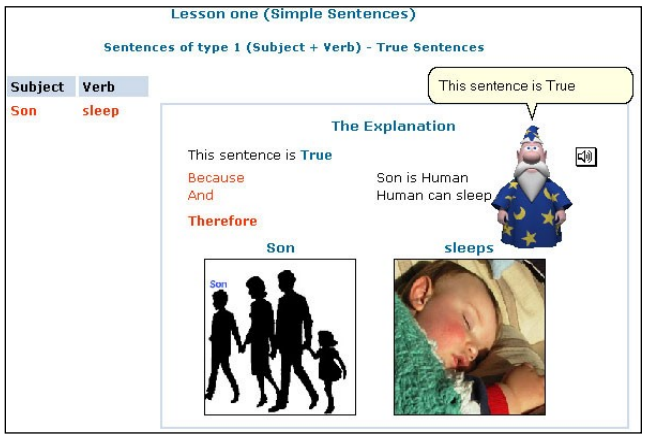



Figure 4. SEMANTIC Dictionary System accepts the created Sentence of Type 1.

Another example is that SEMANTIC dictionary system rejects the created sentence and generates possible category strings to be matched with it as illustrated in Figure 5. The words utilized to build a sentence of type 2 are: 'the subject: Woman, the verb: cook, the object: Newspaper'. SEMANTIC dictionary system rejects this sentence because: 'Woman is human', and 'Newspaper is information media' and 'human cannot cook information media', and it creates possible class string which is 'Human can cook Vegetarian Food' and selects an object randomly from this class, such as 'Potato', and therefore, it corrects the semantical meaning of the sentence and becomes: 'Woman cooks Potato'. Same approach that SEMANTIC deal with the wrong semantic meaning for sentences of type 1; see Figure 6. This lesson exposes the intelligent feature of SEMANTIC dictionary system such as accepting or rejecting the created sentences on its semantical meaning, correcting the wrong sentences and generating correct sentences randomly. In addition, it provides the explanation for accepting or rejecting any sentence.

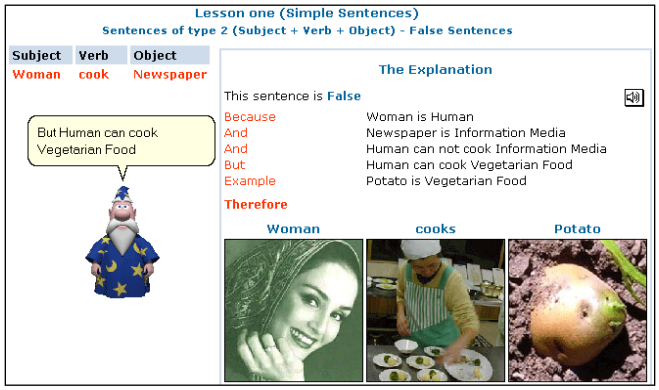



Figure 5. SEMANTIC Dictionary System rejects the created Sentence of Type 2.

 ISBN: 978-1-61208-392-6 

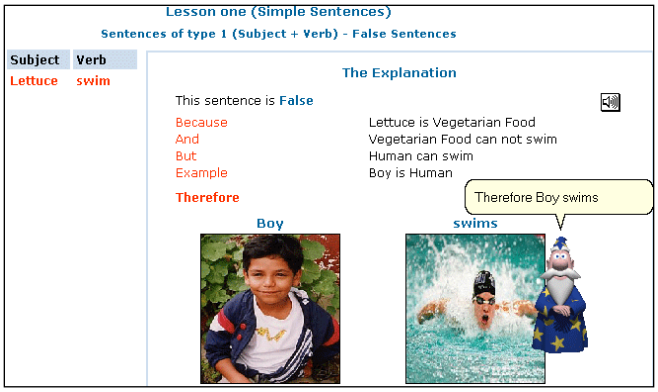



Figure 6. SEMANTIC Dictionary System rejects the created Sentence of Type 1.

**Lesson2** – Singular and Plural Nouns: This lesson teaches the students the comparison between singular and plural nouns in English language and presents the rules that help them to make the plural of the singular nouns. Further, SEMANTIC dictionary system generates four possible sentences from the words Subject (S), Verb (V) and Object (O) that are exposed in the left side, using singular & plural nouns rules i.e. [singular (S) singular (O), singular (S) plural (O), plural (S) singular (O) and plural (S) plural (O)].

For example: The words; Subject (S): Man, Verb (V): Cut and Object (O): Tomato. The sentences that SEMANTIC dictionary system creates are:

- Singular (Subject & Object): Man cuts Tomato
- Plural (Subject) & Singular (Object): Not proper in the real world
- Singular (Subject) & Plural (Object): Man cuts Tomatoes
- Plural (Subject & Object): Men cut Tomatoes

This example indicates that only the second sentence is not proper in the real world, which means that more than one man would not cut one tomato. Therefore, SEMANTIC dictionary system gives a message 'Not proper in the real world'. The intelligent part in this lesson is that SEMANTIC dictionary system generates sentences from three words (Subject, Verb and Object) depending on the singular and plural nouns rules as well as make sure that the generated sentence is proper in the real world. When the generated sentence is not proper in the real world, it shows this message 'Not proper in the real world'.

**Lesson 3** – The Articles: this lesson teaches the students the definitions and types of the articles, as well as teaches the rules for using the articles with nouns in English.

SEMANTIC generates four possible sentences from (Subject, Verb and Object) by using the singular and plural nouns rules, and the articles rules. For example, the sentences that are created from the words (Man, cut and Eggplant), respectively; the sentences that SEMANTIC dictionary generates are:

- Singular (subject) & Singular (object)
  - Indefinite (subject) & Indefinite(object): A Man cuts an Eggplant
  - Definite (subject) & Definite(object):The Man cuts the Eggplant
- Plural (subject) & Plural (object)
  - Indefinite (subject) & Indefinite(object): Some Men cut some Eggplants
  - Definite (subject) & Definite(object): The Men cut the Eggplants

This lesson has some intelligence; for example, SEMANTIC dictionary system selects the appropriate article and produces sentences with grammatical correctness and semantical meaning.

**Checker Tool:** this tool checks the phrase semantics of simple sentences (type 1 and type 2) that entered by the student. In addition, it checks the spelling of words and when any word is misspelt, SEMANTIC will display the possible correct words for it and this will assist students to correct their sentences before checking the semantics meaning of these sentences. For example, the entered words: 'girl, read and cherry' respectively as illustrated in Figure 7. The simple sentence composed of these words is correct grammatically but in the real world meaning (semantics meaning) it is impossible. Therefore, SEMANTIC dictionary system rejects this sentence and displays the reasons which are: 'Girl is human', and 'Cherry is vegetarian food' and 'human cannot read vegetarian food'. SEMANTIC creates possible class string which is 'human can read information media', and selects an object randomly from this class such as 'Newspaper' and therefore it corrects the semantical meaning of the simple sentence and becomes: 'Girl reads Newspaper'. Moreover, it demonstrates the corresponding pictures and presents the Arabic translation.

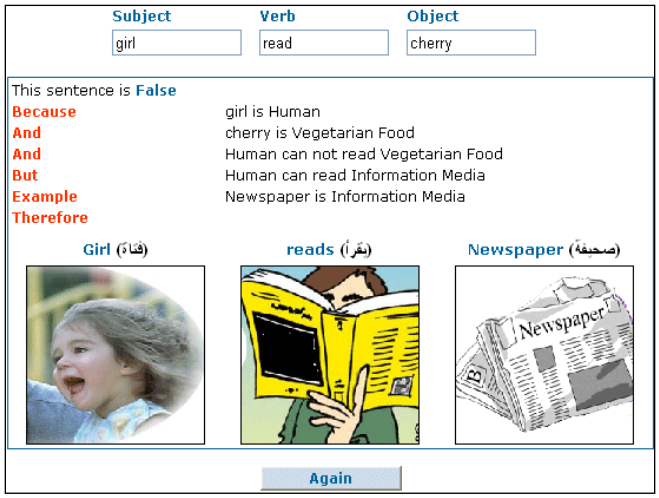



Figure 7. SEMANTIC Rejects the Entered Sentence (General Relationship).

Another example, the entered words: 'man, feed and boy', respectively. SEMANTIC dictionary system accepts this sentence and presents the explanation, as indicated in

Figure 8: 'Man is human', and 'Boy is human', and 'Human can feed Human'. Therefore, 'Man feeds Boy'. It also demonstrates the corresponding pictures and displays the Arabic translation.


Figure 8. SEMANTIC Accepts the Entered Sentence (General Relationship).

## IV. EVALUATING SEMANTICS DICTIONARY SYSTEM

This research presents the Object-Oriented technique as a new approach for the development and evaluation of an Intelligent Multi-Lingual (English - Arabic) dictionary system called SEMANTIC. SEMANTIC dictionary has a number of added features like Tree structures, words classification, simple semantic checks, pictures and voice pronunciation. The SEMANTIC dictionary was evaluated through a large scale case study conducted on elementary schools for girls in the Kingdom of Bahrain to test it for teaching some aspects of English language and to test its level of contribution to students learning. Two groups of a sixth grader (experimental and control) were selected as the subjects of the study. These students were classified into two groups: the first group (Control group) was taught the lessons only under the supervision of a teacher without using the SEMANTIC dictionary system, whereas the second group (Experimental group) was taught the lessons through the SEMANTIC dictionary system under the supervision of a teacher.

Control group: contains 16 students and those who taught SEMANTIC dictionary system tools under the supervision of a teacher without using the SEMANTIC dictionary system.

Experimental group: includes 16 students and those who were taught the same tools as the first group and under the supervision of a teacher but the SEMANTIC dictionary was used.

Sixteen hours of learning time were allotted for the experimental and control groups. The same words and teaching materials (dictionary, lessons, semantic checker, etc.,) were introduced to both groups. The couple of groups did three quizzes that generated by using SEMANTIC dictionary system. The control group answered the quizzes on paper while the experimental group answered it on the computers through the website of SEMANTIC dictionary system.

Finally, students in the experimental group and their English teacher have been asked about their opinions of the SEMANTIC dictionary system.

The achievement test includes three quizzes; quiz 1 (simple sentences), quiz 2 (singular & plural nouns) and quiz 3 (the articles).

Quiz 1 – Simple Sentences: This quiz tests the students' understanding and their ability to create simple sentences. It has two types of questions. Four grades for each question.

Quiz 2 – Singular and Plural Nouns: This quiz examines the students' understanding of singular and plural nouns as well as testing their ability to create sentences. It has two types of questions. Four grades for each question.

Quiz 3 –The Articles: This quiz examines the students' understanding of the articles and singular and plural nouns as well as examines their ability to create sentences. It has one question and eight grades for it.

## V. RESULTS ANALYSIS

The results of students' achievement test of the SEMANTIC dictionary system indicated that the students in the experimental group gained higher grades in the quizzes than the students in the control group.

Figure 9 shows the average grades for the couple of questions of quiz 1 for the students in both groups (experimental and control). This figure indicates that students in the experimental group on average obtained high grades in quiz 1 (simple sentences) especially in the first question. As illustrated in Figure 9, students of the control group on average increased 0.56 on question 2, while students of the experimental group on average decreased by 0.38. However, students of the experimental group on average gained higher grades in this question than students of the control group suggesting that the SEMANTIC dictionary system could help students effectively learn English words and create simple sentences.

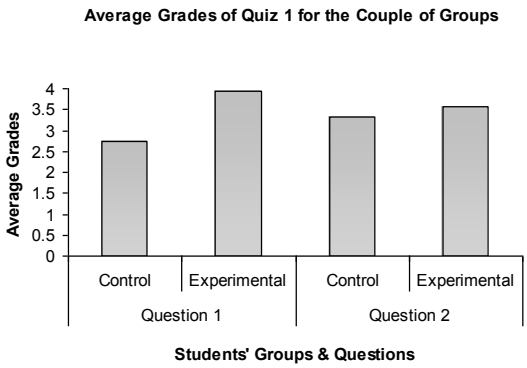



Figure 9. Average Grades of Quiz 1 (Simple Sentences).

Figure 10 exposes the average grades for the couple of questions of quiz 2 for students in both groups (experimental

and control). This figure indicates that students of the experimental group on average acquired high grades in quiz 2 (singular & plural nouns) especially in the first question. As demonstrated in Figure 10, students of the control group on average decreased by 0.06 on question 2 and students of the experimental group on average decreased by 0.69. Even so, students of the experimental group on average achieved higher grades in this question than students of the control group suggesting that the SEMANTIC dictionary system could help students effectively learn singular & plural nouns, and generate sentences that has singular & plural nouns.

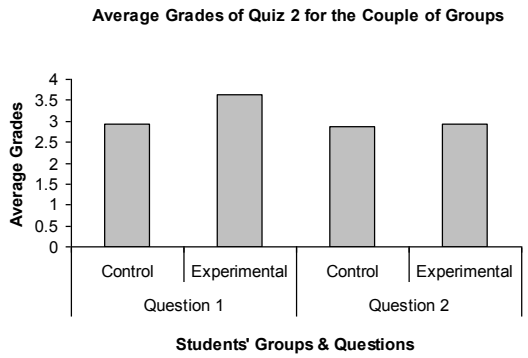



Figure 10. Average Grades of Quiz 2 (Singular & Plural Nouns).

Figure 11 presents the average grades for question 1 of quiz 3 for students in both groups (experimental and control). This figure indicates that students of the experimental group on average gained high grades in quiz 3 (the articles). In this quiz, students of the experimental group obtained on average around 5 and students of the control group on average acquired about 3. These results are low because of the time spent for teaching the two groups "The Articles" lesson score is less than other lessons suggesting that the SEMANTIC dictionary system could help students effectively learn the articles, and organize sentences that contains the articles, as well as the singular & plural nouns.

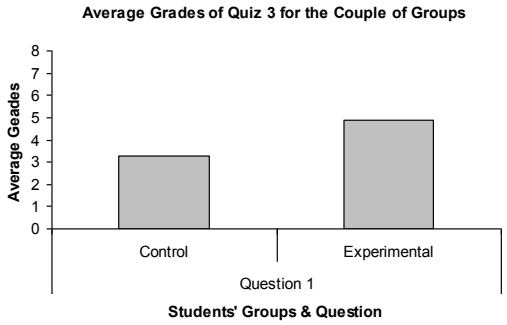



Figure 11. Average Grades of Quiz 3 (The Articles).

## VI. SURVEY RESULTS

The survey of students' opinions towards the SEMANTIC dictionary system deduced that students of the experimental group responded positively towards the SEMANTIC dictionary system; the majority of them enjoyed learning with this system, and would like to learn with this system again. In addition, they suggested publishing this system online because it helps them to learn and understand some aspects of English language easily and in a better way, as well as they can repeat the lessons and testing their understanding any time they want. Moreover, SEMANTIC dictionary system increases their cognitive semantic meaning of words and sentences and assists them to create correct simple sentences grammatically and realistically.

When the English teacher of students have been asked about her opinions towards the SEMANTIC dictionary system (students page), she said that "The SEMANTIC dictionary system is very useful for students especially the beginners and the young learners, because it enables them to understand the meaning of words and at the same time it helps to organize simple sentences and build them with grammatical correctness and semantical meaning. Displaying the classification of the word by Tree structure and covering the words with corresponding pictures makes the process of writing sentences and remembering the meaning easily and helpfully. Also, the pictures attract the students' attention so they go on learning without getting bored. Accordingly, SEMANTIC dictionary system raises their writing skills. Presenting an attractive story with animation and providing voice pronunciation of the words and sentences enhance the reading skills for students, as well assist their acquiring vocabulary."

The above results of achievement test and the survey support the significance and the advantage of the SEMANTIC dictionary system on learning English language. If the SEMANTIC dictionary system is used in elementary schools, both the teaching and learning quality should be improved and students should have greater enjoyment and success in language learning.

## VII. CONCLUSION AND FUTURE WORK

This paper presented a new approach for the development and evaluation of an intelligent Object-Oriented dictionary system called SEMANTIC. The work in this research is the first to effectively use the OO concept and enhance UML to deal with English language semantics for simple phrases. SEMANTIC dictionary system was tested on a 6$^{th}$ class of an elementary school for girls in the Kingdom of Bahrain to test its effectiveness in teaching some aspects of the English language and to test its level of contribution to students learning. The analysis of experimental results is very promising.

Complex structures: sentences consisting of more than one structure or phrase or simple sentence with one verb. This will test the adaptability of sentence to handle multiple structures and how these structures can be integrated semantically, grammatically and structurally as one part with one objective. For example, now students can handle the

 ISBN: 978-1-61208-392-6 

simple sentences *'man reads newspaper'*. However, the sentence '*A man reads the newspaper with drinking a cup of coffee in the afternoon*', consists of a number of structure with complex semantics and inter-relationships. Future work can focus on developing the approach to handle such complexities.

## ACKNOWLEDGMENT

We would like to thank Ahlia University for providing the support and our students and their English teacher for working on developing and testing some of our research ideas.

## REFERENCES

[1] Al-A'ali Mansoor, "Pre-editing and Recursive-Phrase Composites for a Better English-to-Arabic Machine Translation", Journal of Computer Science, Science Publishers, USA, 2007, pp. 1-11.

[2] Al-A'ali Mansoor, "An Efficient Technique for Identifying and Storing Syntactic and Semantic English Natural Language Sentence Patterns", WSEAS Transactions On Systems, vol. 6, issue 1, 2007, pp. 173-178.

[3] Asil Oztekina, Dursun Delenb, Ali Turkyilmazc, and Selim Zaimd, A machine learning-based usability evaluation method for eLearning systems, Decision Support Systems, vol. 56, 2013, pp. 63–73.

[4] Bonnie Piller and Mary Jo Skillings, "English Language Teaching Strategies Used by Primary Teachers in One New Delhi", India School, TESL-EJ, vol. 9, 2005, pp. 1-23.

[5] Haytham S. Al-sarrayrih, Lars Knipping, and Erhard Zorn, "Evaluation of a MOODLE Based Learning Management System Applied at Berlin Institute of Technology Based on ISO-9126", Hasselt, Belgium, 2010, pp. 880-887.

[6] Huiwei Cai, "E-learning and English Teaching. IERI Procedia", vol. 2, 2012, pp. 841-846.

[7] Jaspers, J.G.M., Kanselaar, G. Kok, and W.A.M, "Productive language use with IT'S ENGLISH, Computers & Education", vol 21, issue 1-2, 1993, pp. 67-75.

[8] Mansoor Al-Aali and Samia Yousif, "An Intelligent Multi-Lingual Object Oriented Dictionary System", Proceedings of the 11th WSEAS International Conference on SYSTEMS, Agios Nikolaos, Crete Island, Greece, July 23-25, 2007, pp. 468-473.

[9] Masoud Hashemi and Masoud Azizinezhad, "Teaching English To Children: A Unique, Challenging Experience For Teachers, Effective Teaching Ideas", Procedia - Social and Behavioral Sciences, vol. 30, 2011, pp. 2083-2087.

[10] Purva Chhabra, "Use of E-Learning tools inteachingEnglish", International Journal of Computing & Business Research ISSN (Online): 2229-6166, 2012.

[11] Sally Wu, "Knowledge based approaches for English-Chinese interpretation teaching and learning", Procedia Social and Behavioral Sciences, vol. 2, 2010, pp. 2383-2387.

[12] Samia Yousif and Mansoor Al-A'ali, "Object Oriented Techniques for an Intelligent Multi-Purpose English Language Dictionary System", 9th ICEIS – NLPCS 2007, Funchal-Madeira/Portugal, 2007, pp. 246-255.

[13] Shehdeh Fareh, "Challenges of teaching English in the Arab world: Why can't EFL programs deliver as expected", Procedia Social and Behavioral Sciences, vol. 2, 2010, pp. 3600–3604.

[14] Shulamit Kotzer and Yossi Elran, "Learning and teaching with Moodle-based E-learning environments, combining learning skills and content in the fields of Math and Science & Technology", Heraklion, Crete-Greece, 2012, pp. 122-131.

[15] Wenli Tsou, Weichung Wang, and Hung-yi Li (). "How computers facilitate English foreign language learners acquire English abstract words", Computers & Education, vol 39, issue 4, 2002, pp. 415-428.

[16] Xueyun Ji, "Analysis and Design for Object-oriented Multi-tier Architecture of Public Opinion Survey System Based on UML", Procedia Engineering, vol. 15, 2011, pp. 5445-5449.

[17] Hassan Gomaa, Designing Concurrent, Distributed, and Real-Time Applications with UML, ISBN: 9780321951816, Jan 2014.

 ISBN: 978-1-61208-392-6 

# A Self-Made Personal Explanation Aid for Museum Visitors

Ayako Ishiyama
Tama Art University, Tokyo Institute of Technology
Tokyo, Japan
ishiyama@tamabi.ac.jp

Satoru Tokuhisa
Keio University
Kanagawa, Japan
dangkang@kmd.keio.ac.jp

Fusako Kusunoki
Tama Art University
Tokyo, Japan
kusunoki@tamabi.ac.jp

Shigenori Inagaki
Kobe University
Hyogo, Japan
inagakis@kobe-u.ac.jp

Takao Terano
Tokyo Institute of Technology
Kanagawa, Japan
terano@dis.titech.ac.jp

*Abstract*—**Explanation of museum exhibits must give useful and adequate information to museum visitors. However, good explanation costs a lot and is hard to be maintained by museum curators. This is the case with voice and/or video guidance, or smartphone applications. To cope with the issue, this paper proposes a novel personal support aid: Stamp-On Developers Toolkit (Stamp-On/DT) for visitors, developed by the visitors themselves. Stamp-On/DT consists of smart devices with explanation contents and 'stamp' devices attached to corresponding exhibits. The unique features of Stamp-On/DT are summarized as follows: (1) the digital contents of the corresponding explanation can be created by both visitors and curators, (2) the contents are easily described with conventional web tools such as HTML, CSS, or Java script, and (3) users are only required to save exhibited images in the same exhibited location with the same names. To validate the effectiveness of Stamp-On/DT system, we have conducted a workshop in a museum to let visitors create digital contents and then we have evaluated their performance.. From the workshop experiments, we conclude that Stamp-On/DT is an effective, easy and interesting aid in understanding museum exhibits.**

***Keywords-*** ***tangible user interface; digital content; museum explanation***

## I. INTRODUCTION

The purpose of museums is to collect, store, and educate people with different exhibits. In recent years, lifelong learning has become active and schools have created comprehensive classes. Therefore, demand for education in museums is increasing. With regard to the opportunity for visitors to learn about museum exhibitions, digital exhibition support systems and experiential exhibitions have increased. We have surveyed to identify the expectation of curators from museum visitors. The participants of the survey have indicated that they hope for visitors to have interests in the exhibits, to observe the exhibits more comprehensively, and to feel familiar with the exhibits. Because most visitors often enjoy video games, museum exhibition support systems are required to be both interesting and enjoyable for visitors so that they can engage in observing the exhibits.

Based on such background, in this paper we propose a novel personal support aid: Stamp-On Developers Toolkit (Stamp-On/DT) for visitors, developed by the visitors themselves. The rest of the paper is organized as follows: In Section II, we present a literature survey to highlight the current problems; In Section III, we describe the system configuration and functions of the proposed system; In Section IV, we explain the usage of Stamp-On/DT; In Sections V and VI, in order to validate the effectiveness of Stamp-On/DT, we carry out workshop experiments, then give the findings and discussions; Finally, Section VII concludes the paper.

## II. LITERATURE REVIEW

### *A. Study on Museum Exhibit Explanations*

There are many studies on digital explanations for museum exhibits aimed at people accustomed to interactive stimuli, such as video games [1][2][3][4][5][6][7][8].

Such digital explanations have the same structure as video games. If visitors stand in front of a given exhibit, the digital explanation starts. There are interactive elements to push buttons for more details, but in general, visitors watch the exhibit passively. Experts on exhibits system developments (system experts) are responsible for creating such digital content. Therefore, to fix and/or modify these digital contents, hard work from system experts is required.

 ISBN: 978-1-61208-392-6 

### B. Authoring Tool for Museum Exhibit Explanation

Koleva et al. [9] developed an authoring tool that curators are able to use to connect 3D digital content and sounds for exhibits with a visual programming language. Even with this tool, however, system experts must prepare the 3D parts in advance. Roussou et al. [10] made a website to be used for museum learning, in which they use the pictures drawn by eleven years old children. In Roussou et al.'s study, they report that children made a paper prototype for the web contents. However, finally, a professional web designer created the actual website. Also they reported that the children's pictures required much time to digitize.

### C. Digital Education Tools in Museums

Many museums, including the British Museum and the Louvre, have a digital presence on the Internet. People can watch exhibits remotely [11][12]. On the other hand, Google created a virtual museum for access on the Web in cooperation with different museums, including the National Museum of Western Art [13]. In addition to the Web, museum–display-support applications such as 'Tohaku Navi' [14] and 'e-Museum' [15] are employed. People can confirm the availability of certain exhibitions before visiting a given museum. Okumoto et al. [16] described that watching images and exhibit commentary on the Web before attending a museum was more effective for visitors than using the museum exhibit support system without watching the online commentary prior to visiting the museum. However, it is difficult for all visitors to learn information about exhibits in advance from a museum website. Furthermore, Okumoto et al. indicated that visitors only watched museum exhibits briefly because visitors were preoccupied with awareness of digital content.

### D. Summary of the Survey and Research Statements

Currently, experts are required to make digital exhibition support systems. If only experts create the content, there is a limitation in that modifying existing content or adding new content requires considerable time. Although digital exhibition commentary has a level of interactivity because visitors can press a button, visitors mostly watch the exhibition support system passively. There is also a limitation in that visitors observe digital content more closely than the actual exhibits. Therefore, we believe that museum support systems require a mechanism that can help visitors interact more actively with museum exhibits.

From the literature survey, in an exhibition support system, the roles of visitors are considered very low. However, we believe digital contents should be generated by visitors themselves. It can be attained if the contents are easily developed and modified. Furthermore, if visitors are familiar with interactive video games, they are able to enjoy such digital contents interactively. In this paper, we would like to validate such visitor behaviors.

## III. SYSTEM CONFIGURATION AND FUNCTIONS OF STAMP-ON/DT

Stamp-ON/DT system is an extension of Stamp-On exhibition support tool [16]. The system configuration and functions are, thus, almost the same we have already reported. Based on the previous paper, we explain the outline.

### A. Hardware

The Stamp-On/DT system hardware is composed of a Nexus 7 tablet, stamp, scanner, special paper, and stationary (Figure. 1).


Figure 1. Stamp-On/DT system hardware.

1) Nexus 7 tablet: We need Chrome browser equipped on Nexus 7 tablet: Chrome. However, devices which satisfy the following conditions also run Stamp-On/DT systems:

*a) Device with a multi-touch screen, which is used to detect four or more point coordinates.*

*b) Browser with JavaScript-compatible software.*

2) The stamp: Aluminium tape is pasted on a stamp from the bottom of the stamp to the side of the stamp. The stamp has dot patterns on the bottom (Figure 2), on which the stamp has four convex points. When visitors press the Nexus 7 tablet with the stamp, the tablet reads the dot patterns of the bottom. Each of the stamp pattern identifies the corresponding information attached on the pattern. The

corresponding digital contents will change through this pattern (Figure. 3).

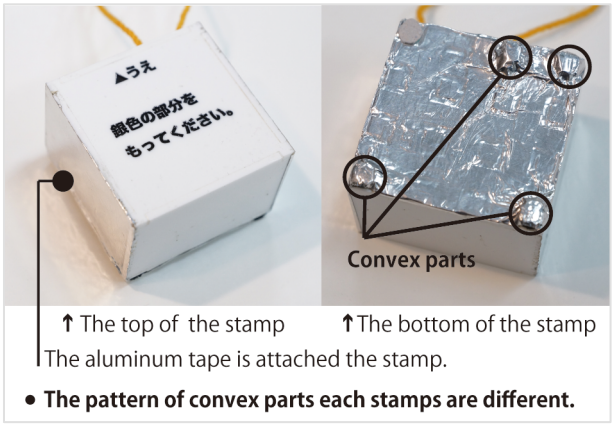



Figure 2. Stamp interface.

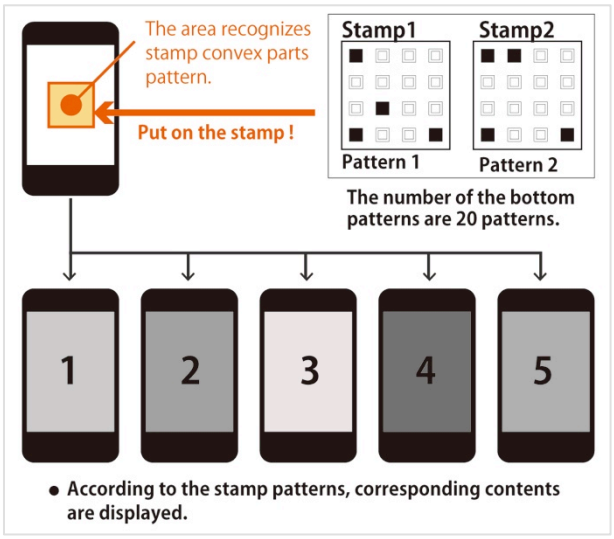



Figure 3. Mechanism to switch digital content.

3) PC and Scanner: A PC and a scanner are required in order to digitize the paper on which visitors write some information on the exhibited items in the form of a single quiz. After the sheet with the quiz is scanned and converted to an image file (jpg format), a support staff will cut unnecessary portions using an image processing software then save the file.

4) Display design sheets and stationeries: To convert digital data, the sheet pre-prints i) a frame in the same screen ratio as the Nexus 7 screen and ii) an area to press the stamp (Figure. 3). Stationeries are used by visitors to write the text and / or to draw the picture to be used.

### B. *Software*

The software used for the proposed system is written in HTML, CSS, and JavaScript. The image file URLs are written in the HTML source file in advance. A new image file is displayed when the image file in the image folder is overwritten. First, visitors bring their paper with the exhibit quiz and commentary to museum staff. Second, the staff overwrites the image file in the specified folder by scanning with the scanner and PC. Finally, the digital content is completed when the staff copies the folder to the Nexus 7 tablet. (If the PC and the Nexus 7 are connected to a web server, the folder is only required to be uploaded.)

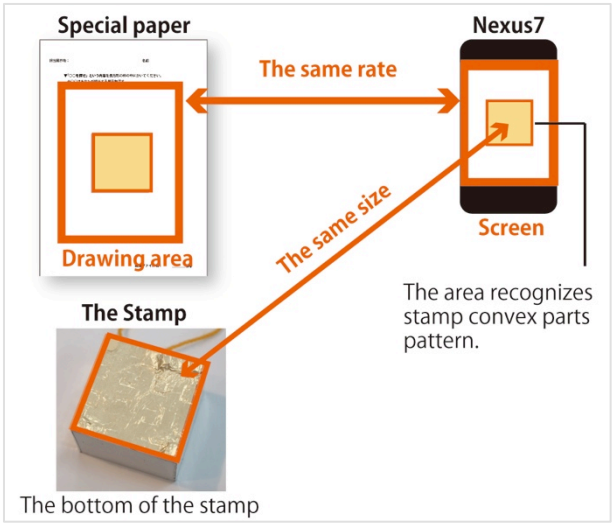



Figure 4. Method for digitizing paper on which visitors draw illustrations for museum exhibits. Method for uploading to Nexus 7 tablet.

## IV. USAGE OF STAMP-ON/DT SYSTEM

Visitors who would like to use Stamp-On/DT are required to perform the following two tasks:

1) To create digital contents (Figure. 5).
2) To play with the digital contents (Figure. 6).

### A. *Task of the Content Creation Phase*

At the first task, visitors are required to follow the steps:

1) Make a quiz regarding a given museum exhibit.
2) Learn about the exhibit while taking notes.
3) Write a quiz related to the exhibition on the sheet with texts and/or illustrations.
4) Scan the sheet then put them into the PC.
5.) Put image files to HTML pages.
6.) Transfer the image files and HTML files to Nexus7.

### B. *Task of the Playing Phase*

At the second task, visitors are required to follow the steps:

1) Place a stamp in front of the museum exhibits.
2) Display the question on the screen of Nexus7.
3) Look for the answer stamp in front of the exhibits.
4) Put the stamp on screen of Nexus7 tablet.
5) Display corresponding contents according to the patterns of the stamp.
6) Display a correct or wrong image. If visitors choose a wrong answer, Nexus 7 displays 'try again'. If visitors choose a correct answer, Nexus7 displays the commentary image which visitors drew.

 ISBN: 978-1-61208-392-6 

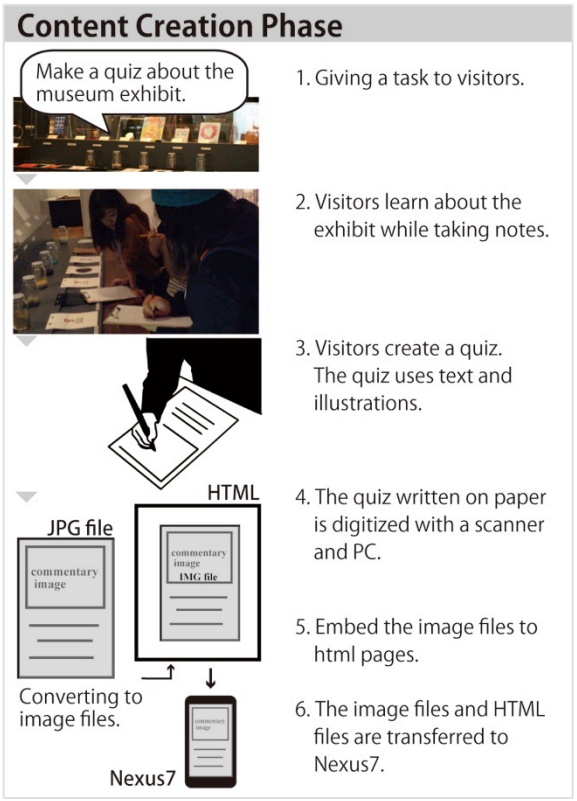



Figure 5. Content Creation Phase.

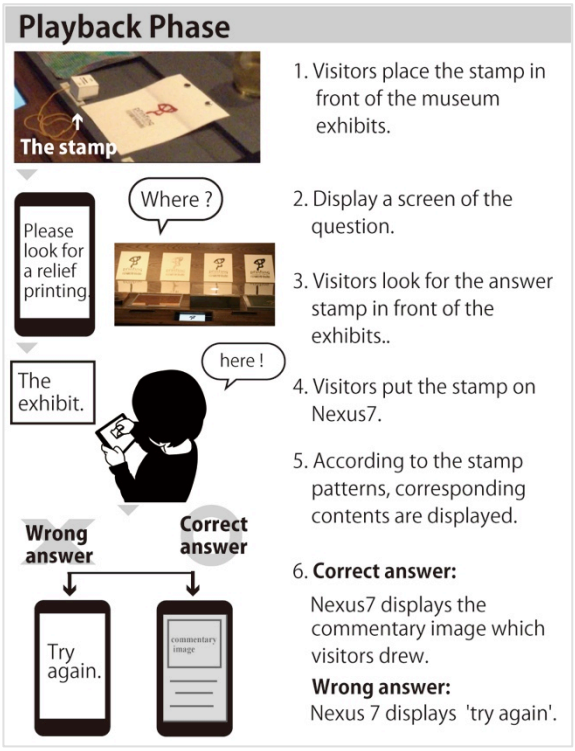



Figure 6. Playback Phase.

## V. EXPERIMENT OF STAMP-ON/DT SYSTEM

### A. Design of WorkShop

To evaluate the effectiveness of Stamp-ON/DT, we have organized a workshop in a museum where visitors were able to observe and enjoy the exhibits actively. Visitors to the workshop were instructed to create digital content to explain the museum exhibits.

When visitors create digital contents, we expect them to show the following behavior:

- Visitors will watch the exhibit more carefully than usual.
- Because visitors are required to create a sheet that explains the exhibit, they need to arrange exhibit information in a header and collect it. Therefore, visitors will understand the exhibit more comprehensively than usual.

### B. Experimental Environment

We conducted an experiment to evaluate our system at the Printing Museum in Tokyo, on Saturday, September 27, 2014. The participants were nine female college students, and none of the participants had seen the exhibits previously.

Three days before the experiment, we trained two students for thirty minutes to assist with the activities of the participants to support digitizing, resizing, and saving the information collected during the experiments. Consequently, on the day of the experiment, the participants had no trouble because of the help provided by the student staff members.

Before the experimental workshop, all participants expressed an interest in printing and enjoyed drawing pictures. We divided the students into two groups (four and five people in each group), and the groups were labeled as Group A and Group B.

For both groups, the required task was to create several quizes regarding the exhibition after observing their assigned exhibits (Figure. 7, Figure. 8). Each person was assigned one of two different exhibits randomly.

After a pre-test, Group A started to create digital contents immediately. On the other hand, after the pre-test, Group B observed the exhibits as usual and required to answer a mid-test. After the mid-test, Group B was required to start to create the corresponding digital contents. As indicated in Table 1, we gave the pre- and post-test to Group A as follows:

- T1. pre-test: the participants answered the test without seeing the exhibits in the museum in advance.
- T3. post-test: the participants answer the test after using the Stamp-On/DT system.

As indicated in Table 6, to Group B, we gave pre-, mid- and post-tests as follows:

- T1. pre-test: the participants answered the test without seeing the exhibits in the museum in advance.
- T2. mid-test: the participants answered the test just after watching the exhibits as regular visitors.

 ISBN: 978-1-61208-392-6 

- T3. post-test: the participants answer the test after using the Stamp-On/DT system.

TABLE I. FLOW OF THE EXPERIMENTS

Group A

| Time | Action | States |
|---|---|---|
| 10min | Pre questionnaire | T1 |
| 1h45min | Digital content creation (Watch Exhibits and production) | |
| 30min | Playing with the digital content they made | |
| 10min | Post questionnaire | T3 |

Group B

| Time | Action | States |
|---|---|---|
| 10min | Pre questionnaire | T1 |
| 20min | Watch as usual | |
| 10min | Intermediate questionnaire | T2 |
| 70min | Digital content creation (Watch Exhibits and production) | |
| 30min | Playing with the digital content they made | |
| 10min | Post questionnaire | T3 |

Figure 7. Experiment participants who observed exhibits.

Figure 8. Subjects drawing picture for exhibit commentary.

Figure 9. Staff digitizing exhibition commentary sheet drawn by participants.

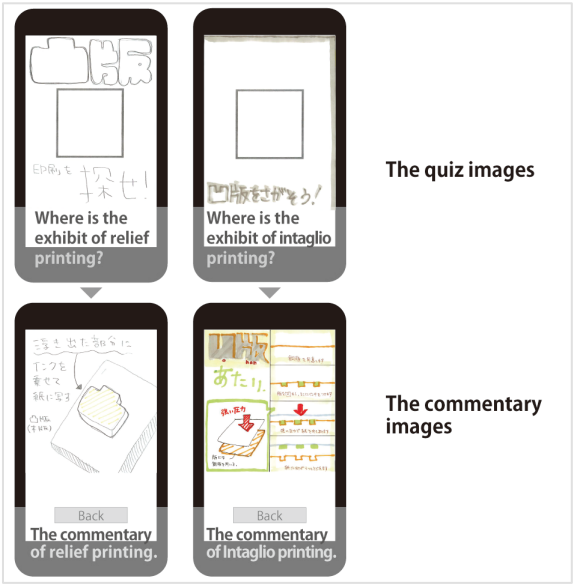



Figure 10. The pictures which subjects painted.

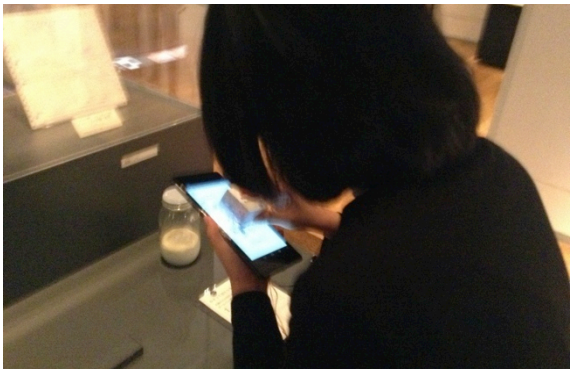


Figure 11. Subjects pressing stamp on Nexus 7 tablet.

## C. *The Objectives of the Evaluations*

We specified the evaluation items of the experiments as follows:

*1) How visitors learnt from the observations on exhibited items.*

- Evaluate the difference in the observations and the learning effects of pre- and post-tests with Groups A and B between (T1) and (T3).
- Evaluate the difference in the observations and the learning effects of pre-, mid-, and post-tests with Grope B among (T1), (T2), and (T3).

*2) How visitors enjoyed the experiences:*

- Evaluate how the visitors enjoyed the proposed systems through the questionnaire analyses.
- Let visitors specify the enjoyable points of the proposed system through questionnaire analyses.

### D. Evaluation Methods

We use the following methods to carry out the evaluation:

*1) Questionnaire Analysis*

*a) Viewing exhibits and learning effects: Multiple-choice and fill-in-the-blank questions were provided in order to determine how the participants learned from the exhibits. Group A answered two questionnaires, before and after the experiment. Group B answered three questionnaires: before, during, and after the experiment.*

For the post-test questionnaire, the participants answered five questions (Q1 to Q5) with five-grade relative estimation.

Q1 and Q2 are related to viewing the exhibits, and Q3, Q4, and Q5 are related to the enjoyability:

Q1. Did you observe the exhibit carefully?

Q2. After the experiment, did you become more careful in observing the general printed information familiar with you and your neighbours?

Q3. Was it interesting for you to make your own descriptions of the exhibited items?

Q4. Was it interesting for you to use the stamp interface?

Q5. Do you like to participate in another similar event, if we would provide the Stamp-On/DT system?

### E. Interview

After the questionnaire sessions, we have made oral interview sessions against randomly selected participants.

*1) About viewing the exhibits: The interview consisted of the following questions: "Did you carefully observe the exhibits?', 'What were different points between your usual museum visits and this experimental observations on the museum exhibits?', 'What were different points between usual explanations of the exhibits and the digital contents you made?'*

*2) About enjoyment: The interview questions were as follows: 'Was it interesting for you to play with Stamp-On?', and 'Was it fun to make your own digital contents?'*

## VI. FINDINGS OF THE EXPERIMENT

### A. Discussion of the Experiments

1) Questionnaire Survey Results.

The answers to the questionnaire survey for Groups A and B are summarized in Table 2. Table 2 depicts experimental results about pre- and post-tests. The sign testing method is applied. The results suggest that there are statistical differences with the 95% reliability. To Group B, we apply the Freedman Testing to pre-, mid-, and post-testing. The results also suggest that there are statistical evidences (Table 3).

Table 4 summarizes the response distributions. Most participants responded positively to all questions. We investigated the response trends after separating the responses obtained from the questionnaire surveys into two groups: positive responses, including 'completely agree' and 'agree', and negative responses, including 'somewhat disagree' and 'completely disagree'. Fisher's exact tests (1×2) showed a statistical significance at 95% level for all items.

TABLE II. THE LEARNING EFFECT ON THE EXHIBIT

**The fill-in-the-blank questions from related to the print.**

| Subjects No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| The pre-test (T1). | 6 | 6 | 4 | 8 | 5 | 4 | 1 | 3 | 5 |
| The post-test (T3). | 15 | 14 | 16 | 18 | 13 | 10 | 8 | 9 | 7 |

p= 0.003906, (p<.05). ↑ The number of correct answers

T1,T3 = states.

**Four questions about the type of printing.**

| Subjects No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| The pre-test (T1). | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 2 |
| The post-test (T3). | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 1 | 2 |

p=0.007812, (p<.05). ↑ The number of correct answers

T1,T3 = states.

TABLE III. TABOUT THE DIFFERENCES BETWEEN PRE-, MID-, AND POST-TESTING

**The fill-in-the-blank questions from related to the print.**

| | The pre-test (T1). | The intermediate test (T2). | The post-test (T3). |
|---|---|---|---|
| SubjectsB1 | 6 | 8 | 15 |
| SubjectsB2 | 6 | 10 | 14 |
| SubjectsB3 | 4 | 6 | 16 |
| SubjectsB4 | 8 | 9 | 18 |
| SubjectsB5 | 5 | 10 | 13 |

Friedman chi-squared = 10, p=0.003906 (p<.05).

T1, T2, T3 = states.

**Four questions about the type of printing.**

| | The pre-test (T1). | The intermediate test (T2). | The post-test (T3). |
|---|---|---|---|
| SubjectsB1 | 1 | 4 | 4 |
| SubjectsB2 | 0 | 4 | 4 |
| SubjectsB3 | 1 | 1 | 4 |
| SubjectsB4 | 0 | 2 | 4 |
| SubjectsB5 | 0 | 4 | 4 |

Friedman chi-squared = 8.375, p=0.01518 (p<.05).

T1, T2, T3 = states.

 ISBN: 978-1-61208-392-6 

TABLE IV. RESULT OF OBSERVATION AND ENJOYABILITY

| About Observation. | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|
| Q1. Do you think you carefully viewed the exhibit? ** | 3 | 5 | 1 | 0 | 0 |
| Q2. Do you think that you tried to more carefully view a printed item that is more familiar to you? ** | 3 | 6 | 0 | 0 | 0 |

**p<.05, 5=completely agree, 1=completely disagree.

| About enjoyability. | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|
| Q3. Do you think that it was fun for you to write a description of the exhibit? ** | 2 | 7 | 0 | 0 | 0 |
| Q4. Did you think that it is fun to press a stamp? ** | 2 | 7 | 0 | 0 | 0 |
| Q5. If there were another event of this type, do you think that you would want to participate? ** | 6 | 3 | 0 | 0 | 0 |

**p<.05, 5=completely agree, 1=completely disagree.

### B. *Results of Interview Survey*

#### 1) *About viewing the exhibits.*

- Participant A: I observed the exhibit carefully more than usual with the intention of preparing a quiz about it.
- Participant B: Commentary must be written to be easy to understand because it will immediately become the corresponding digital contents and will be shown to other participants. I observed the exhibit seriously to try to understand it properly in order to clearly make the contents.

#### 2) *Utterlance of Enjoyable Aspects*

- Participant A: I was impressed at the fact that just after making the quiz, it quickly became the corresponding digital content.
- Participant B: When I pressed the stamp, the immediate reactions the system made was quite interesting.
- Participant B: It was interesting to see the digital contents the other participants developed, because the contents gave me different others' perspectives on their focal points and explanations of the exhibits.

### C. *Summary and Discussion of Experiment Findings*

Based on the questionnaire and interview results, the participants viewed exhibits more carefully with the proposed system than usual visits. All participants suggested that (i) it was pleasant to partake of the interview of the experiment, (ii) creating the digital content is much more interesting than making usual paper contents.

The experimental results have revealed that museum visitors would observe exhibits more carefully than usual visits, if the visitors could create quizzes about the exhibits. Furthermore, all participants have interests in the beautiful printing techniques, which curators of the museum usually use to make explanations of the exhibited items. Therefore, the participants have more interests in the various printings among them in the sense of color, styles, and materials.

When the nine participants used the Stamp-On/DT system at the same time, it was possible for them to produce 18 items of digital contents within 2 hours. These results have shown the superiority of the proposed system against prior digital contents research in the literature [10] on the points of the agility and non-professional support to produce the digital contents.

## VII. CONCLUSIONS AND FUTURE CHALLENGES

This paper has described the design principles, functions, components, usages, and experiments on Stamp-On/DT system, which is a new extension of our Stamp-On [16]. Stamp-On/DT is a toolkit to let museum visitors develop digital contents. The unique features of Stamp-On/DT are summarized as follows: (1) the digital contents of the corresponding explanation can be created by both visitors and curators, (2) the contents are easily described with conventional web tools such as HTML, CSS or Java script, and (3) users are only required to save exhibited images in the same exhibited location with the same names. To validate the effectiveness of Stamp-On/DT system, we have conducted a workshop in a museum to let visitors create digital contents and to have their performance evaluated. From the workshop experiments, we conclude that Stamp-On/DT is an effective, easy and interesting aid in understanding museum exhibits.

From the experimental workshop, we have suggested that i) Stamp-On/DT system is successful to create digital contents in a short time without professional assistances; ii) the participants observed museum exhibits more carefully than usual, and iii) the learning effects on the exhibits observation was also attained.

The proposed system will be further enhanced so that more kinds of tablet devices other than a Nexus 7 can be used in the proposed system. Also, we will prepare manuals and videos, and improve the stamp shapes so that even naïve users can use the stamps.

The other future work includes 1) the improvement of stamp performance, 2) the introduction of the other kinds of hardware devices to assist the usage, and 3) the improvement of manufacture the stamp development.

 ISBN: 978-1-61208-392-6 

## ACKNOWLEDGMENT

We would like to express our gratitude to the curator of the Printing Museum in Tokyo, Mr. Nakanishi, and the staff at the Printing Museum. We would also like to thank the Tama Art University students, who participated in the experiment. This research was supported in part by JSPS, the Grants-in-Aid for Scientific Research (A) (No. 24240100) and the Grant-in-Aid for Challenging Exploratory Research (No. 24650521, 26540186, 24650566 ).

## REFERENCES

[1] Y. Ohashi, H. Mashima, F. Kusunoki, M. Arisawa, "Science Communication from Primary Learning Group to Secondary Learning Group by Adopting Voice Information [in Japanese]," Kagaku Kyoiku kenkyu (Journal of Science Education in Japan) Japan Society for Science Education, 2008, pp. 103-110.

[2] C. Cahill, A. Kuhn, S. Schmoll, W. T. Lo, B. McNallu, and C. Quintana, "Mobile Learning in Museums: How Mobile Supports for Learning Influence Student Behavior," Proceedings of the 10th International Conference on Interaction Design and Children (IDC'11), 2011, pp.21-28.

[3] F. Kusunoki, T. Yamaguchi, T. Nishimura, and M. Sugimoto, "Interactive and enjoyable interface in museum," Proceedings of the 2nd ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'05), 2005, pp.1-8.

[4] D. Raptis, N. Tselios, and N. Avouris, "Context-based design of mobile applications for museums: a survey of existing practices," Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI'05), 2005, pp.153-160.

[5] I. Rose, N. Stash, Y. Wang, and L. Aroyo, "A personalized walk through the museum: The CHIP interactive tour guide," Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI '09), 2009, pp.3317-3322.

[6] T. Yamaguchi, F. Kusunoki, and M. Manabe, "Design of a System for Supporting Interaction in Museums and Zoos with Mixed Media," Journal of Science Education in Japan, 34(2), 2010, pp.97-106. (in Japanese),

[7] K. Yatani, M. Onuma, M. Sugimoto, and F. Kusunoki, "Musex: A System for Supporting Children's Collaborative Learning in a Museum with PDAs," Systems and Computers in Japan, 35(14), 2004, pp. 54–63 (in Japanese).

[8] C.M. Medaglia, A. Perrone, M. De Marsico, G. Di Romano, "A Museum Mobile Game for Children Using QR-Codes," Proceedings of the 8th International Conference on Interaction Design and Children (IDC '09), 2009, pp.282-283.

[9] B. Koleva, S. R. Egglestone, H. Schnädelbach, K. Glover, C. Greenhalgh, T. Rodden and M. Dade-Robertson, "Supporting the Creation of Hybrid Museum Experiences," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09), 2009, pp.1973-1982.

[10] M. Roussou, E. Kavalieratou, M. Doulgeridis, "Children Designers in the Museum: Applying Participatory Design for the Development of an Art Education Program," Proceedings of the 6th international conference on Interaction design and children (IDC 2007), 2007, pp. 77-80.

[11] The British Museum Explore. http://www.britishmuseum.org/explore.aspx (accessed 2014-10-20).

[12] Louvre Online Tours. http://www.louvre.fr/jp/visites-en-ligne (accessed 2014-10-20).

[13] Google Cultural Institute. https://www.google.com/culturalinstitute/home?hl=ja (accessed 2014-10-20)

[14] Tohaku Navi. http://www.tnm.jp/modules/r_free_page/index.php?id=1467 (accessed 2014-10-20).

[15] e-Museum. http://www.tnm.jp/modules/r_free_page/index.php?id=168#e-museum.

[16] M. Okumoto, H. Kato, "The Learning System Linking Pre-Visit and Museum Learning Materials [in Japanese]," Proceedings of Japan Society for Educational Technology 36 (1), 2012, pp.1-8.

[17] A. Ishiyama, F. Kusunoki, R. Egusa, K. Muratsu, S. Inagaki, T. Terano, "Stamp-On: A Mobile Game for Museum Visitors," Proceedings of 6th International Conference on Computer Supported Education (CSEDU 2014), 2014, pp. 200-205.

 ISBN: 978-1-61208-392-6 

# Improving Cache Efficiency of Content-Centric Networking by Using Encoded Addresses

Yuma Kitade
Graduate School of Information Science and Technology, Osaka University
Osaka, Japan
y-kitade@ist.osaka-u.ac.jp

Shingo Ata
Graduate School of Engineering
Osaka City University
Osaka, Japan
ata@info.eng.osaka-cu.ac.jp

Masayuki Murata
Graduate School of Information Science and Technology, Osaka University
Osaka, Japan
murata@ist.osaka-u.ac.jp

***Abstract*—Content-Centric Network (CCN) is expected to become the network architecture in the future for exchanging content without specifying the addresses of nodes. An advantage of adopting CCN is that it improves the availability of network resources by using in-network caching. However, since servers that have a lot of popular content may suffer heavy traffic loads, which lead to frequent updates of caches in neighboring routers, the efficiency of cache usage may be degraded. To solve this problem, we propose a method for widely distributing content by using random encoded addresses. From our simulation results, we show that our method can reduce cache sizes by about up to 75% while still achieving the same cache hit ratio when the access frequencies are biased to a specific server.**

***Keywords–Future Network; Content-Centric Network; Encoded Addresses; Cache Efficiency.***

## I. INTRODUCTION

In recent years, Information-Centric Network (ICN) has been considered as a network architecture for the future, and one implementation of ICN is Content-Centric Network (CCN). Today, communication on the Internet is performed by specifying the identifier of the terminal node (IP address), namely, by node-oriented routing. However, recently, most communications are service-oriented, i.e., end users do not care about which node provides a service or content. It is so usual that end users specify some key words of content/service at search engine like Google, then click the URL returned by the engine. There is thus a gap between how the Internet performs routing and how the Internet is used. CCN is expected to solve this problem by implementing routing that uses information about the content, unifying content-oriented communication.

One of the functionalities of CCN is that it can implement routing without being aware of the location of nodes by routing using information about the content. Performance of the CCN can be improved by creating duplicates of content (a *content cache* hereafter) in any router along the delivery path, and by routing to the nearest node that has the desired content. Furthermore, the content cache can relax from spatial and temporal constraints on content by uncoupling the combination of content and location because content can be arranged to be in any location. That is, the content cache makes it possible to acquire content regardless of where the server with the original content is located and whether the server is running.

To increase the advantages of CCN, efficient creation and placement of content caches is important. Since the content is expected to become more diverse and larger in size in the future, the efficiency of the content cache is essential for effectively finding and using network resources. Therefore, efficient cache algorithms for CCN have been studied. Because research into CCN is still in the early days, evaluations of cache performance are mostly at the level of basic studies. For example, in [1] [2], cache performance at the chunk level in CCN was determined by simulation. However, the network topologies examined in that study were limited to the cascade topology and the tree topology.

In recent years, cache performance has been evaluated in more realistic topologies [3]. In [4], the cache performance in a more general network topology was evaluated. In [5], the cache performance was evaluated by using Mixed-Integer Linear Programming problem. In [6], Bernardini et al. proposed a method for caching only popular content. In [7], a method for enhancing cache robustness was proposed. In [8], a collaborative caching scheme to improve cache performance was proposed. In [9], Rossi et al. conducted a thorough simulation considering network topology, multi-path routing, content popularity, caching decisions and replacement policies. In [10], when individual CCN router had different cache size, the caching performance was evaluated. In [11], Rossini et al. evaluated the influence of multiple servers for the same content.

Although variations in the access frequencies to content are taken into account in many evaluations, such as by modeling with Zipf's law, as exemplified by [9], variations in the access frequencies to servers are often not taken into account. Specifically, the content at a given node is typically assigned randomly. In practice, however, the access frequencies to content servers are heterogeneous, just as access frequencies to content are. In short, the servers (web pages) that attract a lot of access are popular because they have a lot of popular content, and the content that is accessed frequently is concentrated in only some servers. When the locality of content is high, cache updates frequently occur along the peripheral paths of servers that are frequently accessed, and the cache performance is thereby greatly reduced.

The aim of this paper is to take both heterogeneities (not only on access frequency of content but on access frequency of node) into consideration. We propose a network architecture for widely distributing content in CCN to increase the efficiency of content caches. More specifically, this method

 ISBN: 978-1-61208-392-6

does not use content names, but instead uses random encoded addresses that are adapted to the addressing architecture of the network layer protocol for content search and routing. Additionally, content is initially placed in the nodes indicated by the random encoded addresses. We find by simulation that our method improves cache performance.

This paper is organized as follows. First, we give an outline of CCN in Section II, and then propose a method for widely distributing content in Section III. Our simulation results and the effectiveness of our method are given in Section IV, and we present our conclusions and discuss future work in Section V.

## II. Introduction to CCN

Recently many architectures for implementing ICN have been proposed (e.g., DONA [12], PURSUIT [13], SAIL [14] and COMET [15]). The ICN that we target in our work is based on CCN/NDN [16], which is being studied mainly in the US communities.

Communication in CCN consists of *Interest* packets requesting content and *Data* packets supplying content. Thus, whereas content is referenced by a uniform resource locator (URL) in an IP network, it is referenced by hierarchical content names in CCN. For example, an image of an apple can be assigned a name such as /picture/fruit/apple.png.

CCN routers implement forwarding control by content name, and contain three data structures: the content store (CS), pending interest table (PIT), and forwarding interest base (FIB). These are referenced and updated whenever Interest packets arrive. First, the CS is consulted based on the content name given in an Interest packet. The CS contains the list of content names about the contents cached in the router. If the CS contains the content that an Interest packet requires, the router returns the cached content to the client. If not, the FIB and PIT are used to forward Interest and Data packets respectively. The FIB contains information about the next hop for reaching the target content as well as Internet routing tables. CCN routers search the FIB by using the content names in Interest packets, and forward the Interest packets to the next router based on information about the next hop found in the FIB. It also adds and updates the information about the next router in the PIT in order for the Data packets corresponding to the Interest packet to be delivered to the client correctly.

## III. CCN Architecture for Content Distribution

In this section, we propose a CCN architecture for implementing more efficient content distribution. In this architecture, random encoded addresses are first assigned to content, and then the content is placed and Interest packets are routed by using these random encoded addresses.

### A. CCN by Using Encoded Addresses

Figure 1 shows an outline of the proposed CCN that uses encoded addresses. In the proposed architecture, routing is implemented by mapping content names into the address space used for routing in the network layer (e.g., IPv6 or IPv4) instead of using the content names themselves as the target addresses of routing. We call this process “encoding,” and call addresses that have been mapped into the address space used for routing in the network layer “encoded addresses.”

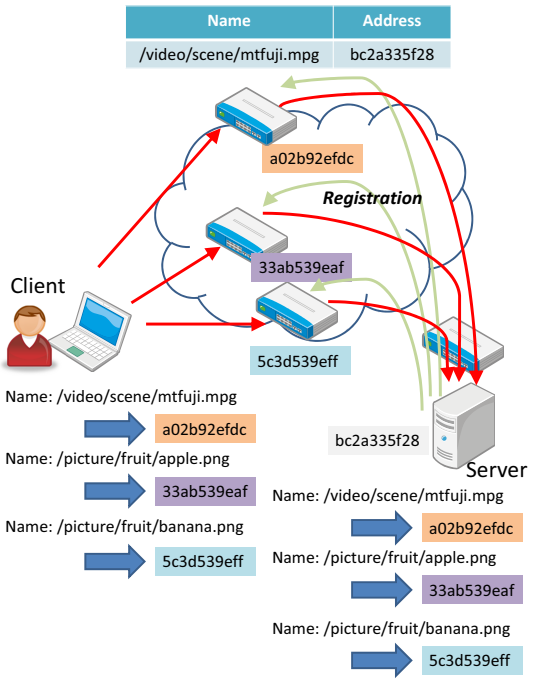



Figure 1. Outline of CCN using encoded addresses

The advantages of routing using encoded addresses instead of content names include (1) the ability to implement CCN without making significant changes to the existing network layer; and (2) the ability to gain additional value from the method of generating encoded addresses. Point (1) means that, if we use IPv4 or IPv6, which are already used on the Internet as the lower layer, and use 128-bit IPv6 addresses as encoded addresses, then we can take advantage of IPv6 routing functions for packet forwarding. Taking full advantage of existing technology is expected to lead to earlier transition to CCN. Point (2) means that, for example, we can use an encoding method to manage the placement and routing of content to be able to widely distribute content by using randomly encoded addresses. The improvement in cache performance by distributing content as described in this paper is achieved by distribution using randomly encoded addresses. More information is given in Section III-B. Also, we note that our architecture has an advantage that the strategy of cache placement is dependent on how to generate the encoded address from the specified content. In other words, we can easily change the strategy by only changing the function of address encoding in the network.

As an example, consider a network layer protocol with an address length of 40 bits, as shown in Figure 1. Now consider the acquisition of a video file named /video/scene/mtfuji.mpg. First, a server that has the object content (a content server) notifies the network that it has the content. We call this “Registration.”

Registration is performed with the encoded addresses of the content. That is, the address `a02b92efdc` is the encoded address of the content /video/scene/mtfuji.mpg, and also means that the network layer address of the node which treats the location of the content and receives the Registration message for the content. Therefore, the Registration message that was

 ISBN: 978-1-61208-392-6

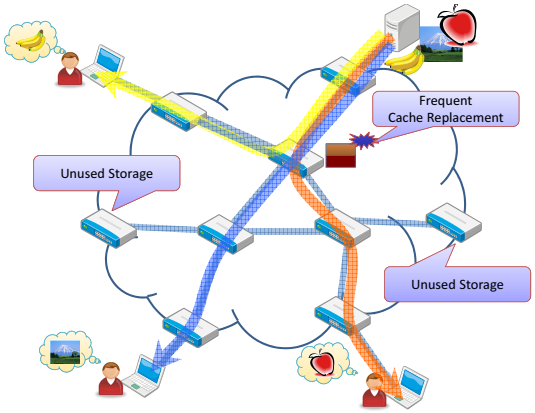



Figure 2. Problem in case of frequently updated caches

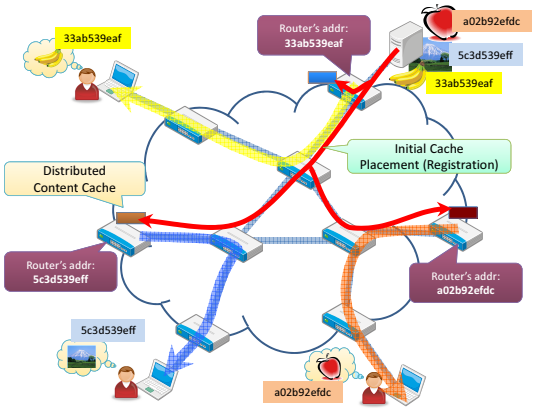



Figure 3. Content distribution with randomly encoded addresses

sent to the address arrives at the node `a02b92efdc` via the network layer protocol. The Registration message contains the content name and the network layer address `bc2a335f28` of the content server. When the node `a02b92efdc` receives the Registration message, it adds the information to a table (i.e., FIB) that maps content names to server addresses (content mapping table).

When a client wants to access the content /video/scene/mtfuji.mpg, it sends an Interest packet to the node that has the encoded address `a02b92efdc` as the network layer address. The Interest packet is sent to the node `a02b92efdc` by the network layer. The node that receives the Interest packet then searches the content mapping table for the content name contained in the message. If the content name is found, the node gets the address of the corresponding content server and forwards the Interest packet to the server. When the content server receives the Interest packet, it sends the content to the client.

### B. Content Distribution by Using Random Encoded Addresses

When content that is accessed highly frequently is concentrated on a specific server, the cache updates occur frequently in neighboring nodes, and the efficiency with which the cache is used deteriorates (Figure 2). To solve this problem, we propose a method for scattering the nodes where content is initially placed by using randomly encoded addresses.

We use a random encoding that offers highly random encoded addresses for content names as a method that readily achieves the above-mentioned content dispersion. For example, this random encoding could employ a method of selecting encoded addresses based on a random hash value calculated from the content names by a hash function. Highly random hash functions, such as Secure Hash Algorithm 1 (SHA-1), are the most suitable for the random encoding.

The randomly encoded addresses also retain their meaning as addresses in the lower routing layer. Moreover, since the addresses are dispersed randomly, a random node can be selected independently of the network topology. Conventionally, it has been necessary to know the topological structure to achieve uniform dispersion of content. However, our method is able to obtain the address of a random node by using randomly encoded addresses, and this enables highly dispersed placement of content. Furthermore, we can also support a locality of content access. If encoded addresses are fully randomized by the name of content, contents would be widely distributed as a whole. On the other hand, if a content is preferable to distribute regionally, we can use a random function with preserving some length of prefix.

In the CCN shown in Subsection III-A, the server that has the content (the content server) does not notify (via a Registration message) the network of the existence of the content, but instead initially places the content in the node that has the address `a02b92efdc`, and the content isn't replaced by other contents. This makes it possible not only for the client to acquire content directly without passing through redundant paths, but also to distribute the content. Figure 3 shows an outline of the CCN for content distribution.

## IV. Performance Evaluation

In this section, we compare the case where content is distributed in a network by using randomly encoded addresses with the case where popular content is concentrated on a specific server without randomly encoded addresses. In this paper, caching is assumed to be performed in chunks, which are portions of the content, rather than in entire content units. Communication is also performed in chunks.

### A. Simulation Scenarios

Performance was evaluated by computer simulation. We use a modified version of ccnSim [9] as the simulator. Specifically, although ccnSim considers the distribution of popularity of content, it also assumes that this content is evenly dispersed across nodes in the network. Thus, because it does not assume that popular content is concentrated on a specific server, which is the subject of this paper, we revised the content placement algorithm in ccnSim so that popular content would be concentrated on a specific server. For comparison, we evaluated cases in which 50% and 70% of content demand is concentrated on a specific server; we also evaluated the case where content is dispersed by using randomly encoded addresses.

The distribution of popular content is implemented by assigning the frequency of demand for each piece of content according to Zipf's law, as follows:

$$f_r = \frac{c}{r^\alpha}, \quad (1)$$

 ISBN: 978-1-61208-392-6

where $f$ is the number of times content is requested, $r$ is the order of popularity of all content, $c$ is a constant, and $\alpha$ is a tuning parameter.

The distribution of popular content greatly depends on the parameter $\alpha$. In papers that evaluate cache performance in CCN [17] [18] [19], values such as 0.8 and 1.0 are used. In addition, the value of $\alpha$ in DailyMotion is about 0.88 [20]. In this paper, we use 0.9 for the value of $\alpha$.

We assume that Interest packet generation is modeled by a Poisson distribution with 100 requests/s, that the frequency of requests for each piece of content obeys Zipf's law, and we also suppose that the nodes do not request content which they themselves own. If the content acquisition for a recent request has not yet completed, the same request is not made again.

Intermediate nodes always cache Data chunks they receive that are not already contained in their caches. Chunks are discarded according to the Least Recently Used (LRU) scheme in the event that the cache becomes full.

After the simulation results had become stable enough, we evaluated our simulation. That is, we calculated the performance metrics after the caches of all nodes had become full and the cache hit ratios had converged. The simulation was finished after the cache hit ratio had converged and a certain period of time had elapsed. Simulations were performed five times while changing the seed of the pseudorandom number generator, and the mean evaluation results were considered.

We assume that the average number of chunks is 100. We varied the cache size of the nodes (the number of Data chunks each node can cache) while comparing the case where content is distributed in the network by using randomly encoded addresses with cases where a specific server contains popular content without randomly encoded addresses. Furthermore, to examine the influence of the number of pieces of content, we performed simulation of the cases of $10^5$ and $10^6$ pieces of content.

A Level 3 Network Topology consisting of 46 nodes [21] [9] was used as the evaluation topology. However, because this topology, shown in Figure 4, consists of only the core network, we used the topology with three end nodes connected for each core node to include the case where end nodes are connected. As a result, the number of nodes is 184 and the graph diameter is 6. The end nodes consist of both servers and clients. The shortest paths by number of hops are calculated beforehand, and the paths between pairs of nodes do not change during the simulation.

Since the efficient utilization of caches is the main topic of this paper, all nodes are assumed to have caches. We also assume that link capacity is sufficient to prevent congestion.

Hereinafter, we refer to nodes that primarily contain content placed using randomly encoded addresses as *repositories*. Our method distributed content across the nodes in a network, and the cache size of each node to have decreased by 10%.

### B. Performance Metrics

We use the cache hit ratio and the hop reduction ratio as performance metrics. The cache hit ratio is the probability that the desired content exists in a node on the path to a repository (or server). Among the nodes $1, 2, \ldots, n$, the node $n$ is the repository (or server), and the number of cache hits in node $i$ is $H_i$. The cache hit ratio $Q$ is then given by (2). In addition, nodes that generate Interest packets first check their own cache for a hit. Interest packets are assumed to hit in node $n$ when they arrive at the repository (or server).

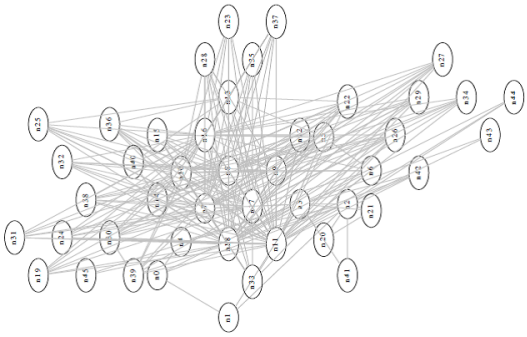


Figure 4. Level 3 topology

$$Q = \frac{\sum_{i=1}^{n-1} H_i}{\sum_{i=1}^{n} H_i} \tag{2}$$

The hop reduction ratio is the mean of the value obtained by dividing the number of hops $d$ through which a Data packet passed by the smallest number of hops $P$ between the node that generated the Interest packet and the repository (or server) that contains the desired content in cases where the node that generated the Interest packet received the Data packet it requested. Nodes are numbered $1, 2, \ldots, n$. The smallest number of hops between node $i$ and node $j$ is $H_{i,j}$. The number of cache hits that occurred in node $l$ for Interest packets generated by node $k$ requesting content $c$ is $D_{c,k,l}$. The repository (or server) of content $c$ is $r_c$. The total number of Interest packets requesting content $c$ generated by node $i$ is $I_{c,i}$. The total number of Interest packets requesting content $c$ is $I_c$. The total number of Interest packets $I$ is thus calculated as follows.

$$I_{c,i} = \sum_{l} D_{c,i,l} \tag{3}$$

$$I_c = \sum_{i} I_{c,i} \tag{4}$$

$$I = \sum_{c} I_c \tag{5}$$

The hop reduction ratio of Interest packets generated by node $i$ requesting content $c$ that have a cache hit in node $j$ is given by (6).

$$p_{c,i,j} = \frac{H_{i,j}}{H_{i,r_c}} \tag{6}$$

The average hop reduction ratio of node $i$ generating Interest packets requesting content $c$ is then given by (7).

$$p_{c,i} = \frac{\sum_{j} p_{c,i,j} D_{c,i,j}}{I_{c,i}} \tag{7}$$

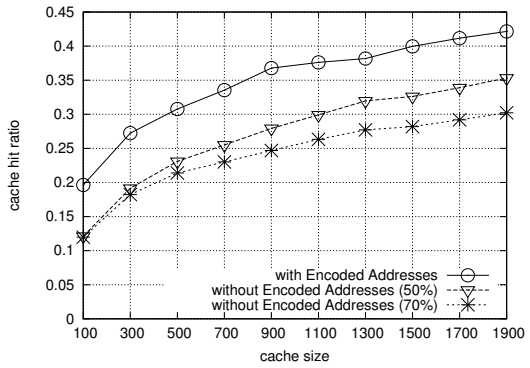



Figure 5. Cache hit ratio with and without randomly encoded addresses ($C = 10^6$)

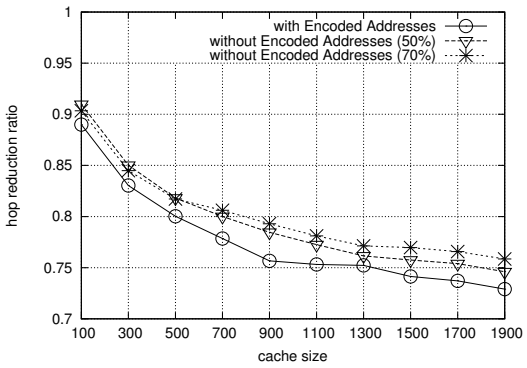



Figure 6. Hop reduction ratio with and without randomly encoded addresses ($C = 10^6$)

The average hop reduction ratio for content $c$ is thus given by (8).

$$p_c = \frac{\sum_i p_{c,i}}{N} \tag{8}$$

The hop reduction ratio is given by (9). In addition, $C$ is the number of pieces of content.

$$P = \frac{\sum_c p_c}{C} \tag{9}$$

Expanding (9) gives (10).

$$P = \frac{I}{NC} \sum_c \sum_i \sum_j \frac{H_{i,j}}{H_{i,r_c} I_{c,i}} \tag{10}$$

The smaller this value, the shorter the response time.

### C. Simulation Results

In Figures 5 - 9, the percentage of content requests that are concentrated on a specific server in the "without Encoded Addresses" cases are indicated in the legend. That is, the label 70% indicates that 70% of all requests are concentrated on a specific server. In addition, cache sizes are given in units of content. That is, a cache size 2000 means each node has capacity for 2000 pieces of content. Hereafter, the cache size in units of content is $S$, and the number of pieces of content is $C$.

From Figure 5 and 6, it is clear that the cache hit ratio and hop reduction ratio improve when randomly encoded addresses are used. This is because the caches of frequently accessed nodes and their neighboring nodes are not updated frequently because the content was placed in random nodes beforehand by using randomly encoded addresses. Specifically, our method of random encoded addresses improves the cache hit ratio by a maximum 9% over the case where 50% of content requests are concentrated on a specific server and by a maximum of 12% over the case where 70% of content requests are concentrated on a specific server. Moreover, randomly encoded addresses can reduce the required cache size significantly by achieving the same cache hit ratio. For example in Figure 5, the cache hit ratio is 0.3, for the case when the cache size is 500 and encoded addresses is used, which is almost the same when the cache size is 1900 without encoded addresses (70%). In other words, encoded addresses can reduce the cache size to about 1/4 to achieve the same performance on cache hit ratio.

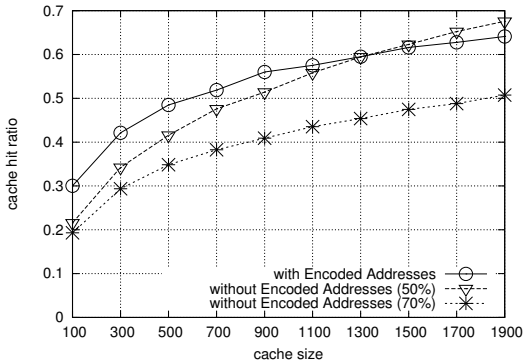



Figure 7. Cache hit ratio with and without randomly encoded addresses ($C = 10^5$)

The hop reduction ratio is improved by a maximum of 3% over the case where 50% of content requests are concentrated on a specific server and by a maximum of 4% over the case where 70% of content requests are concentrated on a specific server. In [17], a scale-free topology of the same order as we used was evaluated, and cache hit ratio and hop reduction ratios were both improved by about 3%. We therefore consider our method to offer significant performance improvement.

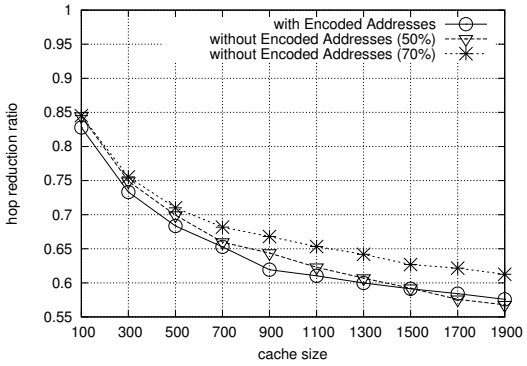



Figure 8. Hop reduction ratio with and without randomly encoded addresses ($C = 10^5$)

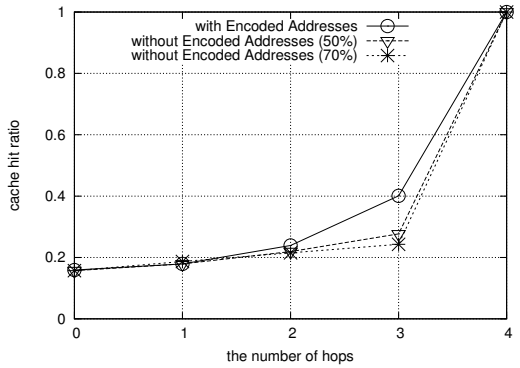



Figure 9. Cache hit ratio every number of hops from clients with and without randomly encoded addresses ($C = 10^6$, $S = 900$, $h = 4$)

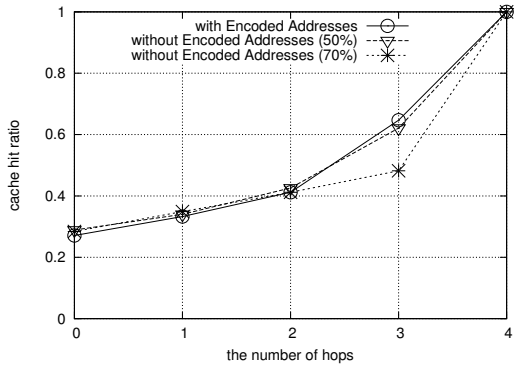



Figure 10. Cache hit ratio every number of hops from clients with and without randomly encoded addresses ($C = 10^5$, $S = 1300$, $h = 4$)

From Figure 7 and 8, for $S = 1300$, our method offers the same performance as the case where 50% of content requests are concentrated on a specific server. To analyze the reasons for this, we assume that the requested content was obtained from the cache of a node on the path to the repository. We therefore calculate the cache hit ratio every number of hops from a client when the smallest number of hops between a client and a repository is $h$. The number of Interest packets sent to repositories is then $R_h$, and the number of times content requested by Interest packets was found in the cache of a node that the number of hops from the client is $i$, is $E_{h,i}$. In addition, when content was acquired from a repository, we treat it as found in the cache of a node where the number of hops from the client is $h$. That is, the number of times the repository is accessed is $E_{h,h}$. The cache hit ratio $C_{h,i}$ from a client to a node $i$ hops away is then given by (11).

$$C_{h,i} = \frac{\sum_{j=0}^{i} E_{h,j}}{R_h} \tag{11}$$

A comparison of the cache hit ratio every number of hops from the client with and without randomly encoded addresses is shown in Figure 9 for the case where $C = 10^6$, $S = 900$, and $h = 4$ (a case where performance is improved by our method), and in Figure 10 for the case where $C = 10^5$, $S = 1300$, and $h = 4$ (a case where performance was not improved by our method). In these figures, the horizontal axis indicates $i$, and the vertical axis indicates $C_{h,i}$.

From Figure 9, it is clear that the cache hit ratios of nodes $h-1$ hops away from the client are greatly improved by using randomly encoded addresses. Although caches in neighboring nodes were frequently updated because popular content was concentrated in a specific server, the content distribution by using randomly encoded addresses resulted in a reduction in the frequency of cache replacements. That is, the problem is solved by our method. The improvement in hop reduction ratio is limited to about 4% because the cache hit ratio up to $h-2$ hops remains almost the same. This is because replacement of caches originally occurred only rarely in nodes near the clients.

Figure 10 shows that the cache hit ratios of nodes $h-1$ hops away from the client are greatly improved by using randomly encoded addresses compared with the case where 70% of content requests are concentrated on a specific server, but the cache hit ratio is not significantly improved compared with the case where 50% of content requests are concentrated on a specific server. With $C = 10^5$, and $\alpha = 0.9$, the most popular 1300 pieces of content account for 50% of content requests. That is, when 50% of content requests are concentrated on a specific server, that server contains the 1300 most popular pieces of content. Therefore, for $S = 1300$, because the neighboring nodes can cache most of the content of the server, their caches are infrequently updated. Therefore, our method hardly improved performance in that case.

Consequently, if CCN routers can cache most of the content that is frequently accessed on a server, the caches in nodes near the servers updated infrequently. As a result, our method offers virtually no performance improvement in those cases. However, we expect that the cache size in CCN routers is not sufficient to cache most of the content that is frequently accessed on a server (such as a YouTube server). Therefore, we conclude that our method is useful.

## V. Conclusions and Future Work

In this paper, we proposed and evaluated a method for widely distributing content by using randomly encoded addresses in a CCN. Using our method makes it easy to disperse content across a network, and improves cache performance. Furthermore, since our method performs routing by the existing network layer protocol, it is expected to lead to an earlier transition to CCN.

In this paper, we assume that a node that has a random hash value calculated from the content names by a hash function always exists. We need to think a specific method of selecting encoded addresses based on a random hash value (e.g., a case where a node that has calculated encoded addresses doesn't exist).

## References

[1] G. Carofiglio, M. Gallo, L. Muscariello, and D. Perino, "Modeling data transfer in content-centric networking," in Proceedings of 23rd International Teletraffic Congress (ITC 23), Sep. 2011, pp. 111–118.

[2] G. Carofiglio, M. Gallo, and L. Muscariello, "Bandwidth and storage sharing performance in information centric networking," in Proceedings of ACM SIGCOMM Workshop on Information-Centric Networking (ICN-2011), Aug. 2011, pp. 26–31.

[3] G. Tyson et al., "A trace-driven analysis of caching in content-centric networks," in Proceedings of ICCCN 2012, Jul. 2012, pp. 1–7.

 ISBN: 978-1-61208-392-6 

[4] E. J. Rosensweig, J. Kurose, and D. Towsley, “Approximate models for general cache networks,” in Proceedings of IEEE INFOCOM 2010, Mar. 2010, pp. 1–9.

[5] S. Wang, J. Bi, and J. Wu, “On performance of cache policy in information-centric networking,” in Proceedings of ICCCN 2012, Jul. 2012, pp. 1–7.

[6] C. Bernardini, T. Silverston, and O. Festor, “Cache management strategy for ccn based on content popularity,” in Proceedings of 7th International Conference on Autonomous Infrastructure, Management and Security, Jun. 2013, pp. 92–95.

[7] M. Xie, I. Widjaja, and H. Wang, “Enhancing cache robustness for content-centric networking,” in Proceedings of INFOCOM 2012, Mar. 2012, pp. 2426–2434.

[8] S. Wang, J. Bi, and J. Wu, “Collaborative caching based on hash-routing for information-centric networking,” in Proceedings of ACM SIGCOMM 2013, Aug. 2013, pp. 535–536.

[9] D. Rossi and G. Rossini, “Caching performance of content centric networks under multi-path routing (and more),” Telecom ParisTech, Tech. Rep., Jul. 2011.

[10] ——, “On sizing CCN content stores by exploiting topological information,” in Proceedings of IEEE NOMEN2012, Mar. 2012, pp. 280–285.

[11] G. Rossini and D. Rossi, “A dive into the caching performance of content centric networking,” in Proceedings of IEEE CAMAD, Sep. 2012, pp. 105–109.

[12] T. Koponen et al., “A data-oriented (and beyond) network architecture,” in Proceedings of ACM SIGCOMM 2007, Aug. 2007, pp. 181–192.

[13] “PURSUIT,” http://www.fp7-pursuit.eu/ [retrieved: 02, 2015].

[14] “Scalable and Adaptive Internet Solutions (SAIL),” http://www.sail-project.eu/ [retrieved: 02, 2015].

[15] “COntent Mediator architecture for content-aware nETworks (COMET),” http://www.comet-project.org/ [retrieved: 02, 2015].

[16] V. Jacobson et al., “Networking named content,” in Proceedings of ACM CoNEXT 2009, Dec. 2009, pp. 1–12.

[17] W. K. Chai, D. He, I. Psaras, and G. Pavlou, “Cache “Less for More” in Information-centric Networks,” in Proceedings of Networking 2012, May 2012, pp. 27–40.

[18] I. Psaras, W. K. Chai, and G. Pavlou, “Probabilistic in-network caching for information-centric networks,” in Proceedings of ICN’12, Aug. 2012, pp. 55–60.

[19] C. Fricker, P. Robert, J. Roberts, and N. Sbihi, “Impact of traffic mix on caching performance in a content-centric network,” in Proceedings of IEEE NOMEN2012, Mar. 2012, pp. 310–315.

[20] Y. Carlinet, B. Kauffmann, P. Olivier, and A. Simonian, “Trace-based analysis for caching multimedia services,” Orange labs, Tech. Rep., 2011.

[21] N. Spring, R. Mahajan, and D. Wetherall, “Measuring ISP topologies with rocketfuel,” in Proceedings of ACM SIGCOMM 2002, Aug. 2002, pp. 133–145.

 ISBN: 978-1-61208-392-6 