

Contour Points based P2P Algorithm for Shape Matching and Image Retrieval

Jianhui Zhao^{1,3,*}, Mingming Ji¹, Mingui Sun², Wenyan Jia², Zhiyong Yuan¹ and Shuping Xiong⁴

¹ School of Computer, Wuhan University, Wuhan, Hubei, 430072, PR China

² Department of Neurosurgery, University of Pittsburgh, PA 15213, USA

³ Suzhou Institute of Wuhan University, Suzhou, Jiangsu, 215123, PR China

⁴ School of Design and Human Engineering, Ulsan National Institute of Science and Technology, Ulsan, 689-798, South Korea

Received: 21 May. 2013, Revised: 4 Sep. 2013, Accepted: 5 Sep. 2013

Published online: 1 Jan. 2014

Abstract: Based on the extracted contours of objects from images, one point to point (P2P) method is proposed for shape matching and image retrieval. Taking contour of one object as reference, contour of another object is transformed to reach their best match, during which the similarity is evaluated by comparison of two corresponding sets of contour points. Translation, scaling and rotation are all considered in transformation, thus the algorithm is robust to objects with different position, size and posture. The experimental results are presented and compared with those from two popular shape based techniques, Hu invariant moments and Zernike moments. Performance of our new approach has proved its efficiency in both matching accuracy and computational expense, and it can be used in related applications together with the other kinds of shape features or even color features, texture features, etc.

Keywords: Contour Points, Point to Point, Shape Matching, Image Retrieval

1 Introduction

The image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some approaches of adding the metadata such as captioning, keywords, or descriptions to images so that retrieval can be performed over the annotation words. Avoiding the use of textual descriptions, another different approach of content-based image retrieval (CBIR) applies kinds of computer vision techniques to image retrieval, aims at retrieving images based on the similarities in their contents (colors, shapes, textures, or any other information that can be derived from the image itself) to a user-supplied query image or user-specified image features.

CBIR has been studied since the early 1990s, with which images are retrieved by the contents, e.g. pixels, features or semantics [1,2,3]. Until now most research works focus on the features of image, e.g. color, shape, texture, and so on. The color based approaches [4,5,6] utilize color features including color histogram, color set, color moment, color coherence vector, color correlogram,

etc. The texture based methods [7,8,9] employ texture features including the gray level co-occurrence matrix, wavelet transform, Markov random field, local binary pattern, etc. The shape based techniques [10,11,12] adopt shape features including boundary chain code, Fourier descriptor, shape moments, etc. For the shape moments related methods, both Hu invariant moments and Zernike moments are used very popularly.

According to the geometric invariant theory, seven classical moment invariants are constructed by Hu [13,14,15], and they are computed from the central moments through order three, while independent to image scale, translation and rotation. The advantage of Hu invariant moments is that the computation speed is very high, but the disadvantage is that the classification accuracy is low, mainly due to it considers only moments with low order, thus details of the objects in image are not well described.

The Zernike moments proposed by Khotanzad and Hong [16,17,18] is based on the Zernike polynomials orthogonal function, which forms a complete orthogonal basis set defined on the unit disc of $x^2 + y^2 \leq 1$. The feature vector of the low-order moments describes overall shape of the target in image, while the feature vector of

* Corresponding author e-mail: jianhuizhao@whu.edu.cn

higher-order moments describes the details of the object. Zernike moments are a set of orthogonal moments with rotation invariant features, and the moments can also be constructed with arbitrarily high orders. Therefore they have the ability to represent more image information and thus can obtain the higher matching accuracy. Of course, more expensive computational cost is the accompanying problem of Zernike method.

Considering that for some applications of shape based image retrieval, contours of the objects are easy to be extracted, no matter directly from the images or from the already segmented regions with other methods. In this paper, one point-to-point (P2P) algorithm is proposed based on the sampled points from object contour, and is used to compare the similarity of different objects by contours. The novel approach can be applied for shape based matching and then for image retrieval, by itself or together with the other kinds of shape features, or even with color features, texture features, etc.

The rest of our paper is organized as follows: the pre-processing techniques for contour extraction are presented in Section 2, details of the P2P method are proposed in Section 3, some experimental results are displayed and analyzed in Section 4, and finally the conclusion is given in Section 5.

2 Contour Points Extraction

Before contour points based shape matching, the image is pre-processed to extract the object contour in it. Since our method is based on the obtained object contour and thus contour extraction is not the focus point of our work, we did not pay attention to those state-of-the-arts while very complex algorithms [19,20,21,22] of contour extraction, such as active contour model, level set method, interactive image segmenting, etc. Instead, a simple algorithm has been developed to extract contours from image with the help of some related functions from OpenCV, and the procedure is described as follows.

First, the `cvSmooth()` function is adopted to smooth the input image, then the `cvThreshold()` function is employed to convert the image into its binary format, from which the targets can be distinguished from the background. Second, we search for the object contours from image with the `cvFindContours()` function, and then utilize the `cvDrawContours()` function to fill the region surrounded by the contours. Finally, `cvDilate()` function and `cvErode()` function are subsequently used to help eliminate the noise and refine the results. As shown in Fig.1, the contour of target can be extracted from image: the 1st column is the original input image, the 2nd column is the corresponding binary image, the 3rd column is the filled region of object, and the 4th column is the extracted contour.

The extracted contours will be further processed with our proposed P2P algorithm for shape based comparison and matching. Considering both operational convenience

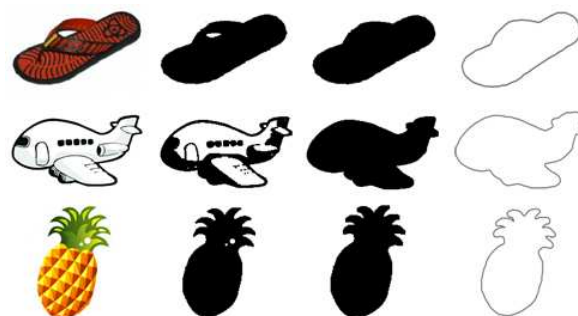


Fig. 1: The extracted contours from images

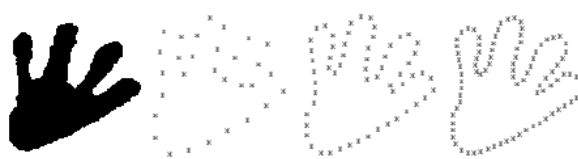


Fig. 2: Contour represented by different numbers of points

and computational expense, a set of sample points are obtained from the object contour. As shown in Fig.2, different numbers of 2D points evenly distributed along the boundary of an object are determined and used to represent the extracted contour. From the left to the right, they are region of object, 30, 60 and 90 points sampled from the extracted contour. Obviously, the contour can be represented in details with more points.

Suppose the set of N points evenly sampled along the extracted contour is defined as P , the i th point of P is represented as $P[i]$, its coordinate vector is $[x_i \ y_i]^T$, and they will be used in the description of our algorithm.

3 Details of P2P Algorithm

The basic idea of our point-to-point algorithm is: taking the contour of one object as reference, transform the contour of another target object to reach their best match based on similarity evaluation through the comparison of two corresponding sets of their sampled points. During transformation, the operations of translation, scaling and rotation are all considered to make sure that the algorithm can deal with or independent with the objects of different positions, sizes and postures.

3.1 Position Translation

For one object, its position can be represented with the center of its shape, while the center point can be computed from the sample points of contour. Based on set P , the center C of one object is calculated as the averaged

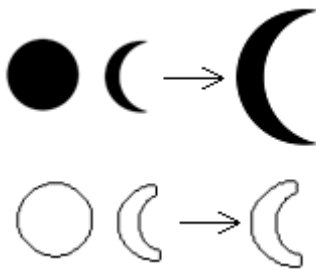


Fig. 3: Size scales based on area (the 1st row) and perimeter (the 2nd row)

value of all x coordinates and all y coordinates from the points of P respectively, by:

$$Cx = \frac{\sum_{i=1}^N(x_i)}{N} \tag{1}$$

$$Cy = \frac{\sum_{i=1}^N(y_i)}{N} \tag{2}$$

Then the center of object is translated to the position of $(0, 0)$, i.e. the origin of the virtual 2D world coordinate system. Since the object is represented through its contour points, the translation is implemented with set P , as:

$$P'[i] = P[i] - [Cx \ Cy]^T \tag{3}$$

where $P'[i]$ is the transformed result of $P[i]$. With position translation, the centers of different objects can become identical, and the P2P algorithm is thus independent to the objects with different locations.

3.2 Size Scaling

Since objects with the same shape may have different sizes, sizes of objects should be measured and normalized before comparison. Area and perimeter are often used to represent the size of one object, and the related scaling results with them are shown in Fig.3. For area based size scaling, the number of black pixels (the 1st row of Fig.3) within the region of object is calculated. After scaling relative to the referenced full moon, size of the crescent moon is increased to have the same area. For perimeter based size scaling, the sum of distances between every two neighbor pixels of contour (the 2nd row of Fig.3) is computed. After scaling relative to the full moon, size of the crescent moon is adjusted to have the same perimeter. Considering that our P2P algorithm works on the sampled points of contour, the perimeter based size scaling is thus adopted since it makes the contours to be compared have the same length.

To speed up the operation of size scaling, computation of all distances between every two neighbor pixels of the contour is replaced with the following method.

The relationships between two neighboring pixels are described as 8 kinds of neighborhoods including north (N), east (E), south (S), west (W), north east (NE), south east (SE), south west (SW), north west (NW). If x coordinates or y coordinates of 2 neighboring pixels are identical, they belong to the group with distance 1 (N, E, S, W); if neither x coordinates nor y coordinates of 2 neighboring pixels are identical, they belong to the group with distance $\sqrt{2}$ (NE, SE, SW, NW). Suppose there are $M1$ and $M2$ pairs of neighboring pixels belonging to these two different groups respectively, size S , i.e. the perimeter of one object is calculated as:

$$S = M1 * 1 + M2 * \sqrt{2} \tag{4}$$

Suppose the reference object is r with size Sr , the target object to be compared is t with size St , target object t can be normalized using the ratio between Sr and St . Through the processing of each pixel in set P , size normalization of object t is achieved by:

$$x'_{i-t} = \frac{Sr}{St} * x_{i-t} \tag{5}$$

$$y'_{i-t} = \frac{Sr}{St} * y_{i-t} \tag{6}$$

where (x_{i-t}, y_{i-t}) is the coordinates of the i th contour pixel of object t , while (x'_{i-t}, y'_{i-t}) is the transformed results of the pixel. With size scaling, reference and target objects have the similar scales, thus P2P algorithm is independent to objects with different sizes.

3.3 Posture Rotation

After the aforementioned position translation and size scaling, the similarity of two different objects can be measured. Taking reference object r and target object t as examples, the difference between them is calculated with their contour points as:

$$D = \sum_{i=1}^N \sqrt{(x_{i-t} - x_{i-r})^2 + (y_{i-t} - y_{i-r})^2} \tag{7}$$

However, the comparison is just between reference r and one posture of target t . To reach the best match among them, target t has to be rotated around its center, and all postures of t are compared with reference r respectively by Formula (7). From the results the minimum difference value is obtained, and thus the posture of t with the best match can be determined. Suppose target t is rotated with angle θ (e.g. 1, 2, ..., 360 degrees, based on the step value of 1.0 degree), contour points of t are transformed through point set P by:

$$\begin{bmatrix} x'_{i-t} \\ y'_{i-t} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_{i-t} \\ y_{i-t} \end{bmatrix} \tag{8}$$



Fig. 4: Matched results of two objects from P2P algorithm

With posture rotation and comparison, the best match between reference object and target object is achieved, which makes the P2P algorithm independent to objects with different postures.

As illustrated in Fig.4, two objects (shoe images in the 1st row) with different positions, sizes and postures are processed with P2P algorithm. After translation, scaling and rotation, the best matched result (contours of two shoes in the 2nd row) is obtained, and the difference value of them is computed with Formula (7).

Of course, the difference value corresponding to the best match can be taken as the similarity between two objects. Through comparing an object with a group of candidates by the similarities, shape based matching and image retrieval can be implemented.

4 Experimental Results

The contour based points to points (P2P) algorithm is first tested on a set of pictures with shoe. To evaluate the performance of being invariant to image translation, scale and rotation, the shoe picture to be matched, its scaled image, and its rotated image are all taken as candidates and put into the set of 35 shoe pictures.

As shown in Fig.5, the first picture in each row is the object to be retrieved, and the other pictures in the same row are 5 of matched results sorted by their similarities with the object, i.e. ordered from the highest similarity to the lower similarity. Results of the 1st row are from Hu invariant moments, results of the 2nd row are from Zernike moments, while results of the 3rd row are from our P2P algorithm. From the experimental results it can be found that our approach obtains the best retrieval performance, even for the candidates with similar shapes but different locations, sizes and postures.

Then the P2P algorithm is tested on 50 toy pictures (Fig.6) and 20 food pictures (Fig.7) consequently. Similar with the experiment on shoe pictures, the object to be matched and the ordered retrieval results are shown in each row from left to right, while the results from Hu invariant moments, Zernike moments and P2P algorithm



Fig. 5: The retrieved results of shoe pictures

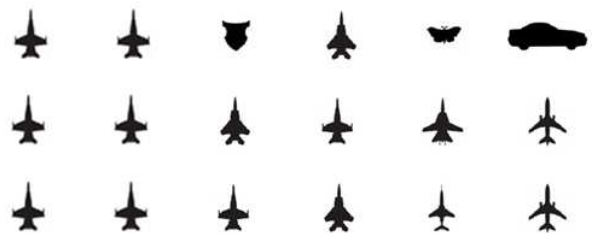


Fig. 6: The retrieved results of toy pictures

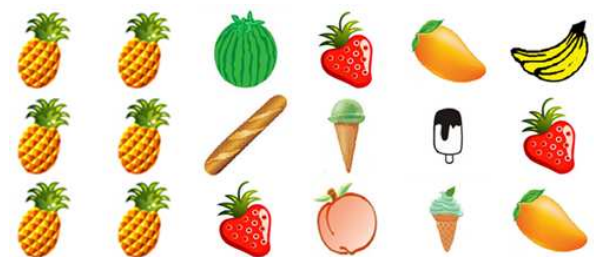


Fig. 7: The retrieved results of food pictures

are shown in the 1st, 2nd, 3rd rows respectively. All the experimental results prove that the proposed P2P method has the best performance for shape matching and image retrieval, while in most cases method of Zernike moments performs better than that of Hu invariant moments.

To test the computing complexity of our P2P based algorithm, the computational costs of Hu method, Zernike method and P2P method on 35 shoe pictures are shown in Fig.8. The horizontal axis represents the number of each picture, while the vertical axis represents the time (unit: ms) spent on both contour points extraction and P2P matching of the related picture. From the comparison, it can be found that computational expense of the proposed P2P method is less than that of Zernike method while more than that of Hu method.

With the help of P2P method, other new algorithms can be implemented through combination of different approaches. For example, Hu method and the proposed P2P method can be combined as: (1) the candidate pictures are matched and retrieved by Hu method, (2) the pictures with lower similarities are ignored, (3) only the

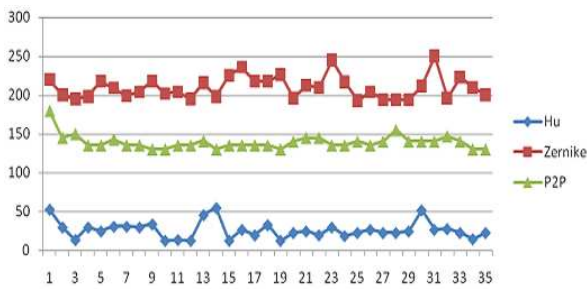


Fig. 8: Computing costs of different methods on shoe pictures

pictures with higher similarities are further processed by P2P method. In this way, the time spent for the whole procedure of only P2P algorithm can be reduced, while the similar objects can still be matched and retrieved from candidates with high accuracy.

The combination of Hu method and P2P method has been implemented and then tested with the set of 35 shoe pictures. The 35 images are firstly retrieved with Hu method, then only the first 10 of the ordered images are further retrieved with P2P method. The computational cost are 7947ms with Zernike method, 5659ms with P2P method, while only 3409ms with the combined approach of Hu method and P2P method.

5 Conclusion

For the pictures with objects that can be segmented easily, the shape based image retrieval is often used, e.g. Hu invariant moments and Zernike moments. Through the contour points of extracted object, a novel point-to-point (P2P) algorithm is proposed in this paper.

In our approach, an object is represented with the set of sample points distributed evenly along its contour, thus the matching of two objects can be implemented by the comparison of two sets of contour points. Taking one object as the reference, the target object is transformed with translation, scaling and rotation to achieve the best match between them, i.e. with the least difference value between their corresponding contour points. During the transformation of object, center computation, perimeter calculation and contour rotation are considered to make our method be independent to the objects with different locations, sizes and postures.

Our algorithm has been tested with some experiments, while both matching accuracy and computational expense are analyzed and compared with the existing Hu method and Zernike method. The presented P2P algorithm has the best performance for both matching and retrieval, and its computing cost is also acceptable. Besides working by itself, our proposed approach can be used together with the other kinds of shape features, or even with existing color features, texture features, etc.

Of course, applications of P2P algorithm are limited due to the fact that the contours of shapes should be determined firstly and properly through applying object segmenting or edge extraction in image. Until now the accurate detections of interesting regions are still very hard to be implemented automatically, and in many cases human interventions are required to instruct the related algorithms or refine the obtained results.

Acknowledgement

This work was supported by National Basic Research Program of China (973 Program No. 2011CB707904), Science and Technology Bureau of Suzhou of China (No. SH201115), Science and Technology Bureau of Wuhan of China (No. 201150124001), Natural Science Fund of Hubei Province of China, R&D Special Fund of Public Welfare Industry of China Meteorological Administration (No. GYHY201106047).

The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

- [1] V. N. Gudivada, V. V. Raghavan, Content based image retrieval systems, *Computer*, **28**, 18-22 (1995).
- [2] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papathomas, P. N. Yianilos, The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments, *IEEE Transactions on Image Processing*, **9**, 20-37 (2000).
- [3] H. Daschiel, M. Datcu, Design and evaluation of human-machine communication for image information mining, *IEEE Transactions on Multimedia*, **7**, 1036-1046 (2005).
- [4] E. Aptoula, S. Lefevre, Morphological Description of Color Images for Content-Based Image Retrieval, *IEEE Transactions on Image Processing*, **18**, 2505-2517 (2009).
- [5] J. Bai, X. Wang, L. Jiao, Image retrieval based on color features integrated with anisotropic directionality, *Journal of Systems Engineering and Electronics*, **21**, 127-133 (2010).
- [6] J. Stottinger, A. Hanbury, N. Sebe, T. Gevers, Sparse Color Interest Points for Image Retrieval and Object Categorization, *IEEE Transactions on Image Processing*, **21**, 2681-2692 (2012).
- [7] P. Sreedevi, W. L. Hwang, S. Lei, An Exemplar-Based Approach for Texture Compaction Synthesis and Retrieval, *IEEE Transactions on Image Processing*, **19**, 1307-1318 (2010).
- [8] R. Kwitt, P. Meerwald, A. Uhl, Efficient Texture Image Retrieval Using Copulas in a Bayesian Framework, *IEEE Transactions on Image Processing*, **20**, 2063-2077 (2011).
- [9] M.H. Rahman, M. R. Pickering, M. R. Frater, D. Kerr, Texture feature extraction method for scale and rotation invariant image retrieval, *Electronics Letters*, **48**, 626-627 (2012).
- [10] S. Biswas, G. Aggarwal, R. Chellappa, An Efficient and Robust Algorithm for Shape Indexing and Retrieval, *IEEE Transactions on Multimedia*, **12**, 372-385 (2010).

- [11] A. Amanatiadis, V. G. Kaburlasos, A. Gasteratos, S.E. Papadakis, Evaluation of shape descriptors for shape-based image retrieval, *IET Image Processing*, **5**, 493-499 (2011).
- [12] M. A. Z. Chahooki, N. M. Charkari, Shape retrieval based on manifold learning by fusion of dissimilarity measures, *IET Image Processing*, **6**, 327-336 (2012).
- [13] M. K. Hu, Visual pattern recognition by moment invariants, *IRE Transactions on Information Theory*, **8**, 179-187 (1962).
- [14] I. Rothe, H. Susse, K. Voss, The method of normalization to determine invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**, 366-376 (1996).
- [15] R. Zhang, L. Wang, An image matching evolutionary algorithm based on Hu invariant moments, *Proceedings of the International Conference on Image Analysis and Signal Processing (IASP)*, Wuhan, China, 113-117 (2011).
- [16] A. Khotanzad, Y. H. Hong, Invariant image recognition by Zernike moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, 489-497 (1990).
- [17] S. Li, M. C. Lee, C. M. Pun, Complex Zernike Moments Features for Shape-Based Image Retrieval, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, **39**, 227-237 (2009).
- [18] Z. Chen, S. K. Sun, A Zernike Moment Phase-Based Descriptor for Local Image Representation and Matching, *IEEE Transactions on Image Processing*, **19**, 205-219 (2010).
- [19] A. K. Mishra, P. W. Fieguth, D. A. Clausi, Decoupled Active Contour (DAC) for Boundary Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 310-324 (2011).
- [20] Y. Liu, Y. Yu, Interactive Image Segmentation Based on Level Sets of Probabilities, *IEEE Transactions on Visualization and Computer Graphics*, **18**, 202-213 (2012).
- [21] L. Zhang, Q. Ji, A Bayesian Network Model for Automatic and Interactive Image Segmentation, *IEEE Transactions on Image Processing*, **20**, 2582-2593 (2011).
- [22] J. Zhao, Z. Zhang, S. Han, C. Qu, Z. Yuan, D. Zhang, SVM Based Forest Fire Detection Using Static and Dynamic Features, *Computer Science and Information Systems*, **8**, 821-841 (2011).



Jianhui Zhao received the PhD degree in Computer Science from the Nanyang Technological University, Singapore in 2004. From 2003 to 2006, he worked as a Research Assistant/Associate in Hong Kong University of Science and Technology. From 2012 to 2013, he worked as a Visiting Scholar in University of Pittsburgh, USA. Currently he is an Associate Professor in Computer School of Wuhan University, China. His research interests include digital image processing, computer vision and computer graphics.



Mingming Ji is a student in Computer School of Wuhan University, China. His research interests include digital image processing and computer vision.



Mingui Sun received the PhD degree in Electrical Engineering from University of Pittsburgh, USA in 1989. He was later appointed to the faculty in the Department of Neurosurgery, and now he is a professor in University of Pittsburgh. His research interests include bio-sensor designs, brain-computer interface, neurophysiological signals and systems, bioelectronics and bioinformatics. He has more than 200 publications.

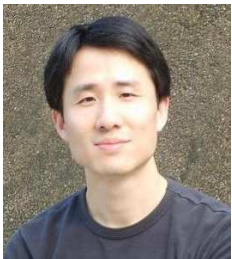


Wenyan Jia received the PhD degree in Biomedical Engineering from Tsinghua University, China in 2005. She is currently a Research Assistant Professor in the Department of Neurosurgery, University of Pittsburgh, USA. Her research interests include brain computer interface and biomedical signal processing.



Zhiyong Yuan is currently a professor at the School of Computer, Wuhan University, P.R. China. His research interests include computer simulation and virtual surgery, image processing and pattern recognition, medical physics, and medical information

system. Zhiyong Yuan received the BS degree in computer application and MS degree in signal and information processing from Wuhan University, in 1986 and 1994, respectively, and a Ph.D. degree in control science and engineering from Huazhong University of Science and Technology in 2008. During 2006-2007, he was a visiting professor at the School of Medicine, University of Pittsburgh, conducting research on computer-based surgical simulation system for training endoscopic neurosurgeons. He has published over 70 papers in related conference proceedings and journals.



Shuping Xiong received the PhD degree in Industrial Engineering (Human Factors Track) from Hong Kong University of Science and Technology in 2008. Currently he is working as an Associate Professor in School of Design and Human Engineering, Ulsan National Institute of Science and

Technology, South Korea. His research interests include human factors, human computer interface.