Recognition for Objects by Relationships Between Attributes

Hiroka Horiguchi*, Kazuo Ikeshiro and Hiroki Imamura

Graduate School of Engineering, Soka University, Hachioji-Shi, Tokyo, Japan

Abstract: Object recognition methods based on attributes have been studied. Conventional methods recognize objects by the presence or absence of attributes. However, the conventional methods have two problems. Firstly, the conventional methods are not able to recognize a target object of which a part of attributes is occluded. Secondly, the conventional methods miss-recognize a target object, which has irrelevant attributes. Therefore, to solve these two problems, we propose the object recognition by relationships between attributes. In this paper, we focus on the face as the recognition object. The proposed method uses relationships as constraints for object recognition using attributes. The proposed method applies two major type constraints. The first constraint is a local constraint, which is applied to a part of attributes. To achieve robust face recognition against occlusion scenes, the proposed method uses the local constraint. And then, the second constraint is a global constraint, which is applied to all attributes. To achieve robust face recognition against irrelevant attributes, the proposed method uses the global constraint. In this paper, to evaluate the effectiveness of the proposed method, we compared the proposed method with a conventional method. We experimented in normal face, occlusion and irrelevant attributes. We used 2580 images of a face which are changed in scale and rotation. Experimental results showed that the recognition ratio of the proposed method is equal to or more than that of the conventional method in normal face, occlusion, and irrelevant attributes.

Keywords: Face recognition, Robust recognition, Relationships between attributes, Occlusion, Irrelevant attributes.

1. INTRODUCTION

In recent years, the object recognition has been required in such as the autonomous cars and image searching. General object recognitions [5-7] are able to recognize a target object by learning the target object previously. Moreover, object recognition methods based on attributes have been proposed [1-3] as more versatile object recognition than general object recognitions recently. For example, in recognition of a unicycle and a bicycle, attributes based object recognition methods describe the unicycle is an object having a wheel and a saddle and describe the bicycle is an object having two wheels and a saddle. Then, object recognition methods based on attributes are able to recognize the unicycle and the bicycle by only learning a wheel and a saddle without learning the unicycle and the bicycle individually because both objects have same attributes. However, conventional methods using attributes have two problems. Firstly, conventional methods are not able to recognize a target object of which a part of attributes is occluded. For example, in a face recognition, in case that conventional methods describe a face as an object having two eyes, a nose, and a mouth, and a mouth was occluded in a scene, conventional methods are not able to recognize a face because a part of attributes lacks. (Figure **1a**) Secondly, conventional methods miss-recognize a target object which has irrelevant attributes. For example, in case that an image has an irrelevant eye and an actual face, although the irrelevant eye and a part of the face are satisfied with the structure of a face nevertheless the combination is irrelevant as a face. (Figure **1b**) Therefore, conventional methods miss-recognize a face by the irrelevant attribute. Therefore, to solve these two problems, we propose the object recognition based on relationships between attributes. Firstly, the proposed

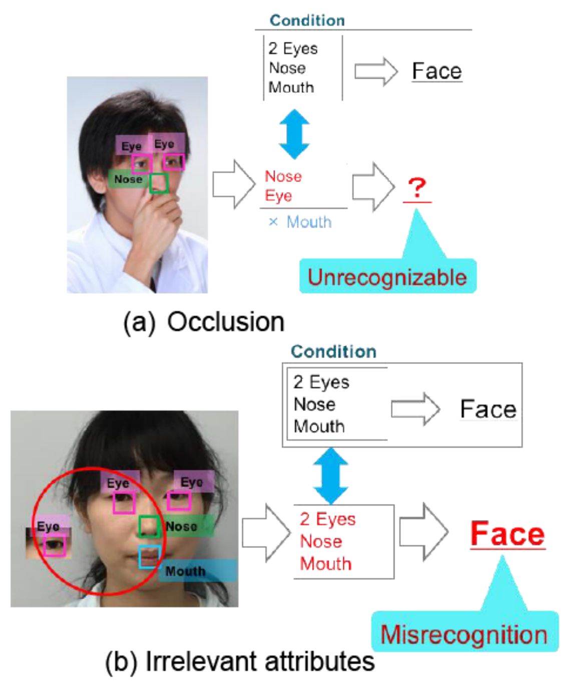


Figure 1: Two problems of the conventional method.

*Address correspondence to this author at the Graduate School of Engineering, Soka University, Hachioji-Shi, Tokyo, Japan;
Tel: 042-691-2211;
E-mail: e16m5217@soka-u.jp

E-ISSN: 2410-2938/16

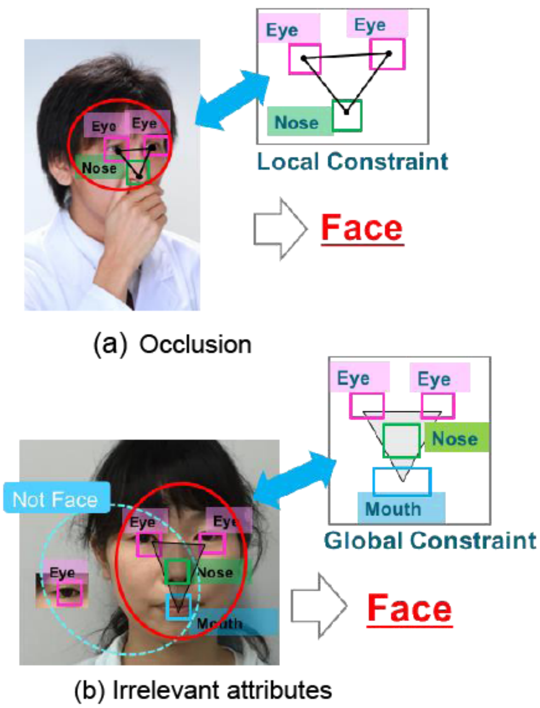


Figure 2: The proposed method.

method is able to recognize a target object of which a part of attributes is occluded because the proposed method describes relationships between other visible attributes of the target object as a local constraint Figure **2a**. Secondly, the proposed method is able to exclude a target object which has irrelevant attributes because the proposed method judges a relevant relationship as an object by using a relationship of all attributes as a global constraint. In this paper, we focus on the facial recognition Figure **2b**.

2. PROPOSED METHOD

2.1. Overview of the Proposed Method

To achieve robust face recognition against occlusion and irrelevant attributes, the proposed method uses relationships between attributes as constraints. In this paper, the proposed method applies two major type constraint. The first constraint is a local constraint which is applied to a part of attributes. And then, the second constraint is a global constraint which is applied to all attributes. To achieve robust face recognition against occlusion scenes, the proposed method uses relationships of angle and length between visible attributes as the local constraint. For example, if a mouth was occluded in a face as shown in Figure **2a**, the proposed method recognizes a face by checking whether angle and length between two eyes and a

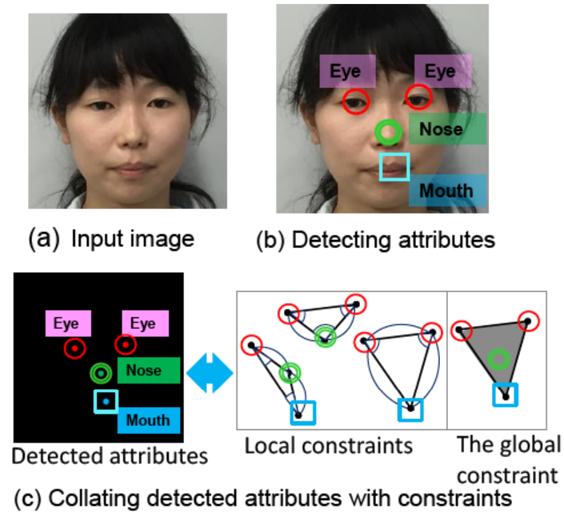


Figure 3: The flow of the proposed method.

nose are suitable as attributes of the face or not. In addition, to achieve robust face recognition against irrelevant attributes, the proposed method uses the relationship of the relative position of all attributes as the global constraint. For example, in case that an eye was miss-detected near a face as shown in Figure **2b**, the proposed method prevents from recognizing the wrong eye as a face attribute by checking whether a nose is surrounded by two eyes and a mouth or not. The local constraint and the global constraint are determined manually. The proposed method recognizes the face when detected attributes are matched any one of the local constraints and the global constraint. Figure **3** shows the flow of the proposed method. Firstly, the proposed method inputs a face image (Figure **3a**). Secondly, the proposed method detects attributes (Figure **3b**). Finally, the proposed method recognizes a face by collating detected attributes with local and global constraints (Figure **3c**). Followings describe a detail of Detecting attributes and collating detected attributes with local and global constraints.

2.2. Detecting Attributes

As shown in Figure **4**, the proposed method detects regions of the eye, the nose, and the mouth by using haar-like feature detection [4]. In this paper, we adopted detectors of a Toolbox for MATLAB [8-10]. These detectors are trained with AdaBoost, Haar-Like feature, and 7000 positive samples for each attribute. We defined the centroids of each attribute as the position of each attribute. And then, the proposed method extracts centroids of these regions as these attribute positions.

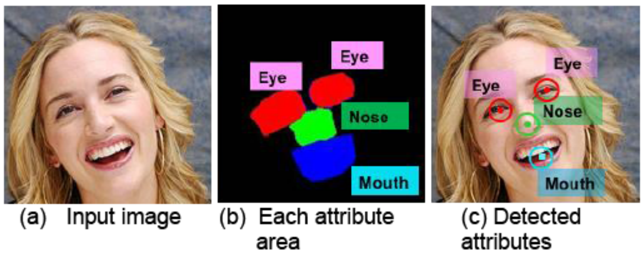


Figure 4: Detecting attributes.

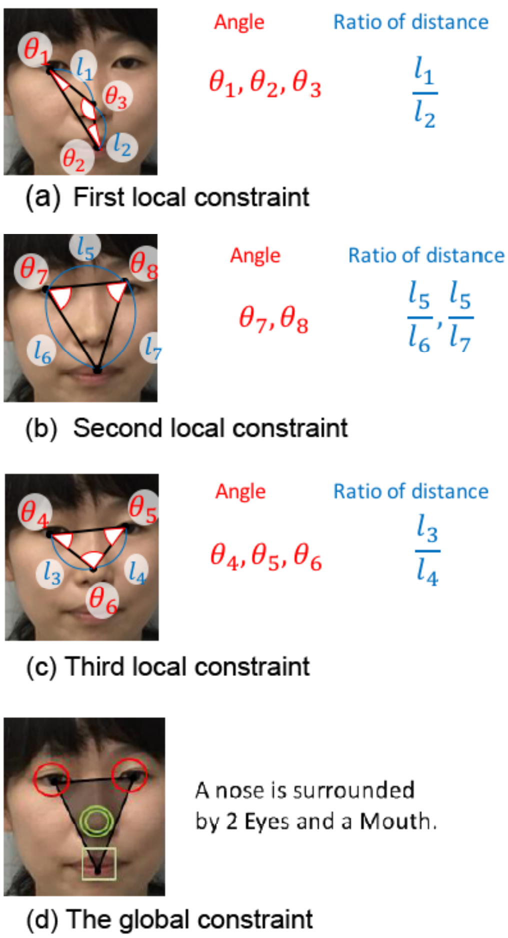


Figure 5: Constraints for the face recognition.

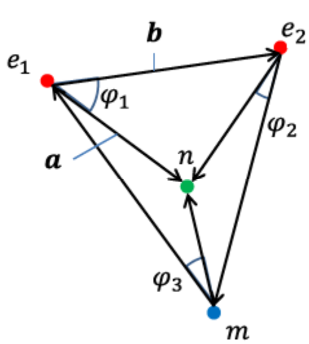


Figure 6: Determination method of the global constraint.

2.3. Collating Detected Attributes with Local Constraints

The local constraint for robust face recognition against occlusion scene consists of three constraints. Each local constraint has parameters of angles and ratio of distances. First local constraint limits angles $\theta_1, \theta_2, \theta_3$, and length ratio $\frac{l_1}{l_2}$ between an eye, a nose and a mouth to thresholds as shown in Figure **5**(**a**). Second local constraint limits angles $\theta_4, \theta_5, \theta_6$, and length ratio $\frac{l_3}{l_4}$ between two eyes and a nose to thresholds as shown in Figure **5**(**b**). Third local constraint limits angles θ_7, θ_8 and length ratio $\frac{l_5}{l_6}, \frac{l_5}{l_7}$ between two eyes and a mouth to thresholds as shown in Figure **5**(**c**).

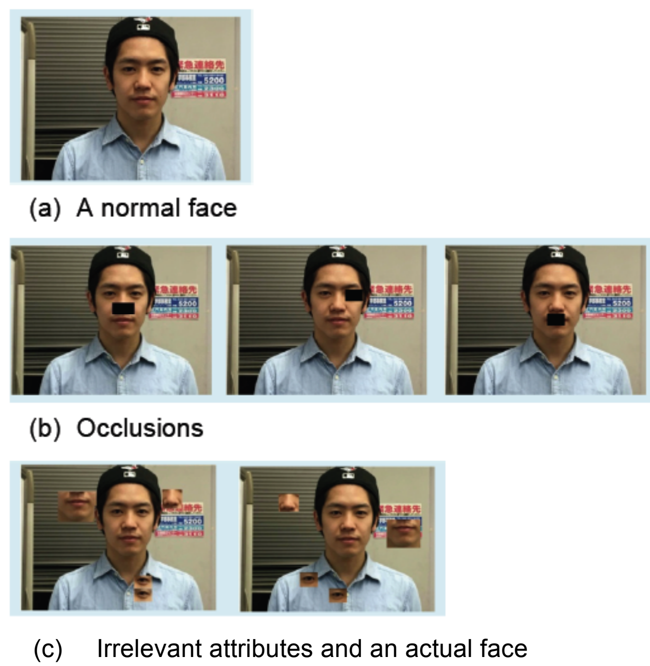


Figure 7: Experimental targets.

2.4. Collating Detected Attributes with the Global Constraint

The global constraint for robust face recognition against irrelevant attributes consists of relative positions of two eyes, a nose and a mouth as shown in Figure **5**(**d**).

The proposed method prevents from miss-recognizing of irrelevant attributes of a face by checking whether a nose is surrounded by two eyes and a mouth or not. As shown in Figure **6**, the proposed method firstly calculates angle φ_1 between vector a from an eye position e_1 to another eye position e_2 and vector b from, e_1 to a nose position n by using

$$\sin\varphi_1 = \frac{a \times b}{|a||b|}. \quad (1)$$

$$\varphi_1 = \sin^{-1}\varphi_1. \quad (2)$$

In the case of φ_2 a vector from another eye position e_2 to a nose position n indicates vector a and a vector from another eye position e_2 to a mouth position m indicates vector b In the case of φ_3 a vector from a mouth position m to a nose position n indicates vector a and a vector from a mouth position m to an eye position e_1 indicates vector b The proposed method calculates angles φ_2 and φ_3 in the same way as φ_1. Then, if $\varphi_1, \varphi_2, \varphi_3$ are more than 0, the proposed method is able to judge that the nose was surrounded by two eyes and the mouth.

(a) Scaling

(b) Rotation of the roll direction

(c) Rotation of the yaw direction

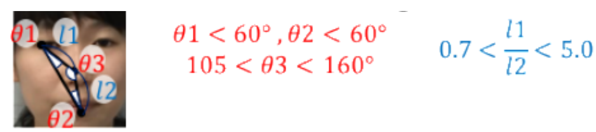
(d) Rotation of the pitch direction

Figure 8: Scale and Rotation change of faces.

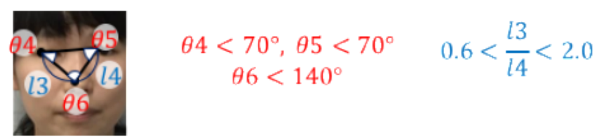
3. EXPERIMENTS

3.1. Experimental Overview

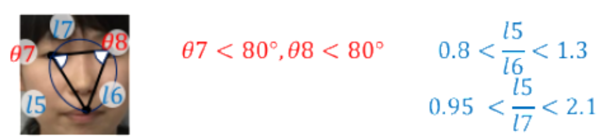
To evaluate the effectiveness of the proposed method, we compared the proposed method and a conventional method [1] in three cases such as (a), (b) and (c) in Figure **7**. The first case is a normal face (Figure **7a**). The second case is the occlusion which any one of an eye, a nose and a mouth is hidden in a face (Figure **7b**). The third case is that irrelevant attributes and actual face are mixed (Figure **7c**). We prepare two types of irrelevant attributes which are changed position for the third case. We used 2580 images of 10 people gathered by ourselves. These images are changed scale and rotation such as followings, for each case.



(a) First local constraint



(b) Second local constraint



(c) Third local constraint

Figure 9: Parameters of local constraints for the experiment.

- Faces which are changed scale from 0.8 times to 1.2 times in step of 0.2 times. (Figure **8a**)
- Faces which are turned roll direction from -90 degrees to 90 degrees in step of 10 degrees. (Figure **8b**)
- Faces which are turned yaw direction from -40 degrees to 40 degrees in step of 10 degrees. (Figure **8c**)
- Faces which are turned pitch direction from -50 degrees to 60 degrees in step of 10 degrees. (Figure **8d**)

In addition, we determined parameters of three local constraints of the proposed method in a trial (Figure **9**).

First local constraint: Maximum angles of θ_1 and θ_2 are able to be 60 degrees. Therefore, thresholds of θ_1 and θ_2 are decided as less than 60 degrees. Minimum angle of θ_3 is able to be 105 degrees, and a maximum angle of it is able to be 160 degrees. Therefore, a threshold of θ_3 is decided as more than 105 degrees and less than 160 degrees. Minimum value of length ratio $\frac{l_1}{l_2}$ is able to be 0.7, and maximum value of it is able to be 5.0. Therefore, a threshold of $\frac{l_1}{l_2}$ is decided as more than 0.7 and less than 5.0 (Figure **9a**).

Second local constraint: Maximum angles of θ_4 and θ_5 are able to be 70 degrees. Therefore, thresholds of θ_4 and θ_5 are decided as lessthan 70 degrees. Maximum angles of θ_6 is able to be 140 degrees. Therefore, thresholds of θ_6 is decided as less than 140 degrees. Minimum value of lengthratio $\frac{l_3}{l_4}$ is able to be 0.6, and maximum value of it is able to be 2.0. Therefore, a threshold of $\frac{l_3}{l_4}$ is decided as more than 0.6 and less than 2.0 (Figure **9b**).

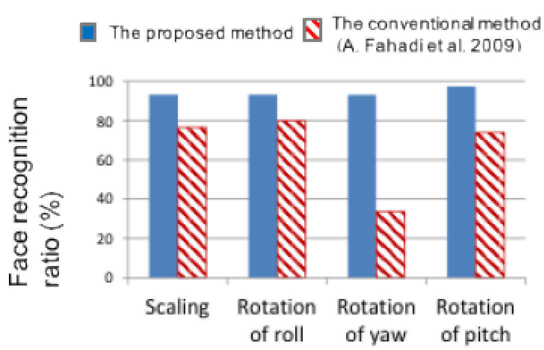


Figure 10: The recognition ratio in the normal face.

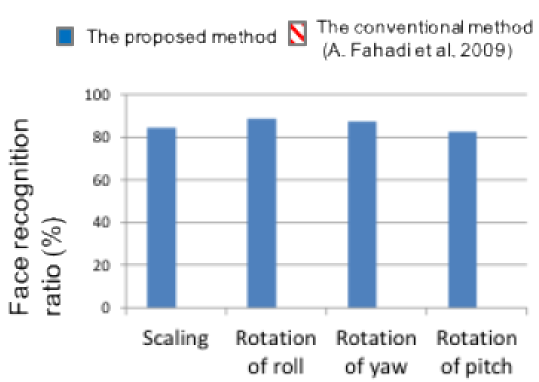


Figure 11: The recognition ratio in occlusion.

Third local constraint: Minimum value of length ratio $\frac{l_5}{l_6}$ is able to be 0.8, and maximum value of it is able to be 1.3. Therefore, a threshold of $\frac{l_5}{l_6}$ is decided as more than 0.8 and less than 1.3. Minimum value of length ratio $\frac{l_5}{l_7}$ is able to be 0.95, and maximum value of it is able to be 2.1. Therefore, a threshold of $\frac{l_5}{l_7}$ is decided as more than 0.95 and less than 2.1 (Figure **9c**).

3.2. Experimental Results

Figures **10**, **11**, and **12** show the recognition ratio of the proposed method and the conventional method in three cases. In the case of normal face, as shown in Figure **10**, the recognition ratio of the proposed method was 93.3% in scaled face, in the face turned in roll direction, and in the face turned in yaw direction and was 97.3% in the face turned in pitch direction. The recognition ratio of the conventional method was 76.7% in scaled face, was 79.9% in the face turned in roll direction, was 33.3% in the face turned in yaw direction and 74.0% in the face turned in pitch direction. In the case of occlusion, as shown in Figure **11**, the recognition ratio of the proposed method was 84.5% in the scaled face, 88.7% in the face turned in roll direction, 87.4% in the face turned in yaw direction and was 82.6% in the face turned in pitch direction. Recognition ratios of the conventional method were 0% in all of faces which are changed scale and rotation. In the case of irrelevant attributes and actual face, as shown in Figure **12**, the recognition ratio of the proposed method was 94.8% in the scaled face, was 82.3% in the face turned in roll direction, was 81.9% in the face turned in yaw direction and was 79.8% in the face turned in pitch direction. The recognition ratio of

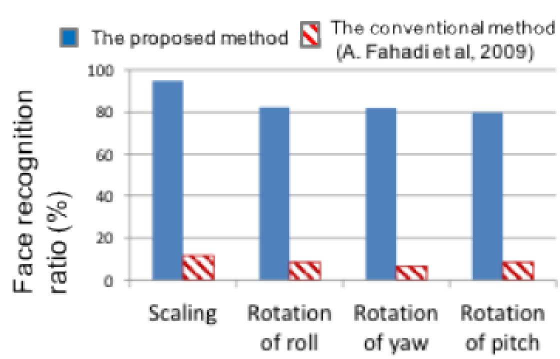


Figure 12: The recognition ratio in irrelevant attributes and an actual face.

the conventional method was 11.7% in the scaled face, was 8.4% in the face turned in roll direction, was 6.7% in in the face turned in yaw direction and was 8.3% in the face turned in pitch direction.

3.3. Discussion

In the case of normal face, as shown in Figure **10**, the recognition ratios of the proposed method more than that of the conventional method in all of faces which are changed scale and rotation because the conventional method sometimes miss-recognized a face by miss-detected attributes and part of attributes of a face, in contrast, the proposed method is able to recognize the face if a relationship between all attributes is relevant. In the case of the occlusion, as shown in Figure **11**, recognition ratios of the conventional method were 0% in all faces which are changed scale and rotation because the conventional method needs two eyes, a nose and a mouth as all attributes of the face to recognize the face, and any one of two eyes, a nose and a mouth lacked in the case of the occlusion. In contrast, the recognition ratio of the proposed method was 84.5% in the scaled face, was 88.7% in the face turned in roll direction, was 87.4% in the face turned in yaw direction and was 82.6% in the face turned in pitch direction. All recognition ratios of the proposed method were high because the proposed method recognized the face by using local constraints when any one of two eyes, a nose and a mouth lacks. In the face which is turned in roll direction, the recognition ratio of the proposed method was 82.6% and was low than other recognition ratios of the proposed method because the proposed method hardly detects a nose when a face turn up, and hardly detects eyes when a face turn down, and the proposed method sometimes miss-recognized attributes as the face when a relationship of miss-detected attributes was accidentally matched any one of constraints. In the case of irrelevant attributes and actual face, as shown in Figure **12**, the recognition ratio of the conventional method was 11.7% in the scaled face, was 8.4% in the face turned in roll direction, was 6.7% in the face turned in yaw direction and was 8.3% in the face turned in pitch direction. These recognition ratios were low because the conventional method miss-recognized the actual face and irrelevant attributes as two faces. The recognition ratio of the proposed method was 94.8% in the scaled face, was 82.3% in the face turned in roll direction, was 81.9% in the face turned in yaw direction and was 79.8% in the face turned in pitch direction. All recognition ratios of the proposed method were high because the proposed method excluded irrelevant attributes by using the global constraint and recognized only the actual face. In the face which is turned in pitch direction, the recognition ratio of the proposed method was 79.8% and was low than other recognition ratios of the proposed method because the proposed method hardly detects a nose when a face turn up, and hardly detects eyes when a face turn down. Therefore, the proposed method easily miss-recognizes a face by irrelevant attributes and part of attribute of actual face. In these three cases, the recognition ratio of the proposed method was more than the conventional method. Therefore, we consider that the proposed method is more effective than that of the conventional method.

CONCLUSION

We have presented a novel method for robust face recognition based on attributes against occlusion and irrelevant attributes. Using relationships between attributes as two major constraints, the proposed method prevents unrecognition due to occlusion and miss-recognition due to irrelevant attributes. Firstly, the proposed method uses local constraints to achieve robust face recognition against occlusion. Secondly, the proposed method uses the global constraint to achieve robust face recognition for irrelevant attributes. Results of experiments about three cases show that the proposed method is more effective than the conventional method. In future works, we experiment using pre-existing dataset [15]. In addition, we work on automatically learning the parameters of relationships between attributes and apply our method to objects other than the face. As an issue of this research, since the proposed method is not able to recognize objects having two or fewer attributes, it is necessary to solve it. We believe that this method can be applied to systems that search images of products having parts and materials required by users.

REFERENCES

[1] Ali F, Ian E, Derek H, David F. Describing objects by their attributes:Computer Vision and Pattern Recognition, 2009. CVPR. IEEE Conference on, Miami 2009; 1778-1785. http://ieeexplore.ieee.org/document/5206772/?denied

[2] Christoph L, Hannes N, Stefan H. Learning to detect unseen object classes by Between-Class Attribute Transfer:Computer Vision and Pattern Recognition, 2009. CVPR. IEEE Conference on, Miami 2009; 951-958. http://ieeexplore.ieee.org/document/5206594/?denied

[3] Tamara LB, Alexander CB, Jonathan S.Automatic Attribute Discovery and Characterization from Noisy Web Data:In Computer Vision, ECCV 2010 - 11th European Conference on Computer Vision, Proceedings, Part I. Berlin 2010; 663-676. http://link.springer.com/chapter/10.1007/978-3-642-15549-9_48

[4] Paul V, Michael J. Rapid Object Detection using a Boosted Cascade of Simple Features:Computer Vision and Pattern Recognition. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001; 1: I-511-I-518. http://ieeexplore.ieee.org/document/990517/?denied

[5] Rainer L, Alexander K, Vadim, P. Empirical analysis of detection cascades of boosted classifiers for rapid object detection:Pattern Recognition: 25th DAGM Symposium, Magdeburg 2003; 297-304. http://link.springer.com/chapter/10.1007/978-3-540-45243-0_39

[6] Yann L, Fu H, Leon, B. Learning methods for generic object recognition with invariance to pose and lighting: Computer Vision and Pattern Recognition. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on 2004; 2: II-97-104 . http://ieeexplore.ieee.org/document/1315150/?denied

[7] Gabriella C , Christopher D, Lixin F, Jutta W, Cédric B. Visual categorization with bags of keypoints:In Workshop on Statistical Learning in Computer Vision, ECCV. Prague 2004; 1-22. http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.72.604

[8] Modesto S, Oscar S, C. Guerra , M. Hernández. ENCARA2: Real-time Detection of Multiple Faces at Different Resolutions in Video Streams: Journal of Visual Communication and Image Representation 2007; 18: 130-140 http://www.sciencedirect.com/science/article/pii/S1047320306000782 https://doi.org/10.1016/j.jvcir.2006.11.004

[9] MC Santana, OD Suárez, DH Sosa, JL Navarro. Using Incremental Principal Component Analysis to Learn a Gender Classifier Automatically: 1st Spanish Workshop on Biometrics 2007,Girona, Spain.

[10] Modesto S, Oscar S, Antón-canalís L. J Lorenzo-navarro. Face and Facial Feature Detection Evaluation: Third International Conference on Computer Vision Theory and Applications, VISAPP08 2008, January. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.77.7484

[11] Guodong G, Li SZ, Kapluk C. Face recognition by support vector machines: Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), Grenoble 2000; 196-201. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=840634&isnumber=18088

[12] Takeshi M, Toshimitsu K, Osamu H. Joint Haar-like features for face detection: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, Beijing, 2005; 2: 1619-1626 . http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1544911&isnumber=32976

[13] P.Jonathon P, Harry W, Jeffery H, Patrick J R. The FERET database and evaluation procedure for face-recognition algorithms: Image and Vision Computing, 1998; 16: 295-306. http://www.sciencedirect.com/science/article/pii/S026288569700070X https://doi.org/10.1016/S0262-8856(97)00070-X

[14] Matthew T, Alex P. Face recognition using eigenfaces: Proceedings 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR Maui 1991; 586-591. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=139758&isnumber=3774

[15] Nicolas G, Daniela H, James LC. Estimating Face Orientation from Robust Detection of Salient Facial Features: Proceedings of Pointing 2004, ICPR, International Workshop on Visual Observation of Deictic Gestures, Cambridge, UK. http://www-prima.inrialpes.fr/perso/Gourier/Faces/HPDatabase.html

Received on 18-11-2016 Accepted on 10-12-2016 Published on 24-12-2016

http://dx.doi.org/10.15379/2410-2938.2016.03.02.01