Modified Deformable Parts Model for Side Profile Facial Feature Detection

Pisal Setthawong and Vajirasak Vanijja

School of Information Technology, King Mongkut's University of Technology Thonburi, Thailand
51500701@st.sit.kmutt.ac.th, vajirasak@sit.kmutt.ac.th

Abstract. Deformable Parts Model(DPM) is a facial feature detection approach. Though the approach is accurate, robust, and works well for a wide range of facial profiles, when faced with a side profile, the typical approach produces less than satisfactory results. This paper discusses about issues faced when attempting to detect facial features on the side profile and proposes modifications to the DPM approach so that it works with detection facial features on side profiles.

Keywords: Deformable Parts Model, Facial Feature Detection, Side Profile Facial Feature Detection, Image Processing.

1 Introduction

Facial Feature Detection is an important image processing task in which important facial features are detected and marked for later use. Systems such as facial recognition systems [1] and certain face pose estimation systems[14] rely on the facial feature system to provide an accurate estimate of key facial features which would provide an ideal starting point for later facial analysis. If the initial facial feature detection process has not provided accurate facial feature positions, the accuracy from the later processing would be severely degraded or simply would not work.

There had been many facial feature detection approaches that have been developed over the years, and one of the popular approaches is Deformable Parts Model (DPM). DPM is not the fastest approach proposed, but the approach provides robust and accurate results in facial feature detection over a wide variety of facial profiles. Though typical DPM approach in facial feature detection can detect many profiles accurately, at certain side profiles, the detection of features is less than ideal. This paper will illustrate why typical DPM approaches for facial feature is not ideal when detecting features for in a side profile, propose suggestions on how to improve the detection rates for side profiles, and provide a side by side comparison between the an open-source implementation of DPM and the proposed modified approach in detecting facial features in side profile.

2 Background

This section explains about the DPM approach in facial feature detection and issues that are present when attempting to detect facial features on side profiles.

B. Papasratorn et al. (Eds.): IAIT 2013, CCIS 409, pp. 212–220, 2013.
© Springer International Publishing Switzerland 2013

2.1 Facial Feature Detection Approaches

One of the first approaches in detecting facial feature is to create independently trained detectors of each individual facial feature [16]. Detectors are trained for the eyes, nose, mouth, face regions, and etc. The AdaBoost based detectors [15] and Haar's Classifier are popular approaches in this area. Though the independent trained detectors can detect facial features, one of the major weakness of this approach is that it detects many false positive facial features. The extreme local nature of the approach is a contributor to the high false positive rate as illustrated in Figure 1. To deal with the high false positive rates, providing a geometric configuration can help lower the rates. The detection is done with the independent individual detectors in which are set as candidate features in which would later be scored based on the geometric bias of the features to select the most likely positions for the features.

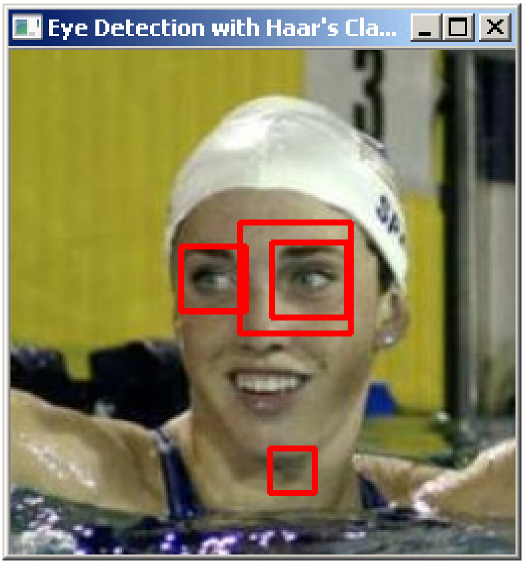

Fig. 1. The detection of the Eye Region with Haar's Classifier Independently usually returns many false positive results

The DPM approach [2] changes from the two stepped independent detectors and geometric configuration into merging both the processes into a single model. The DPM is defined where there features and their set of connections between pairs of features much like an undirected graph where vertices are the features and the edge are the connection between the pairs of features. The DPM detector then estimates the feature positions by using a single scoring function consisting of the local feature model and the deformation cost using an optimization function. Due to the accurate results of DPM, the approach has been used in many successful facial feature detection systems [4,8,10] and one of the results are illustrated in Figure 2.

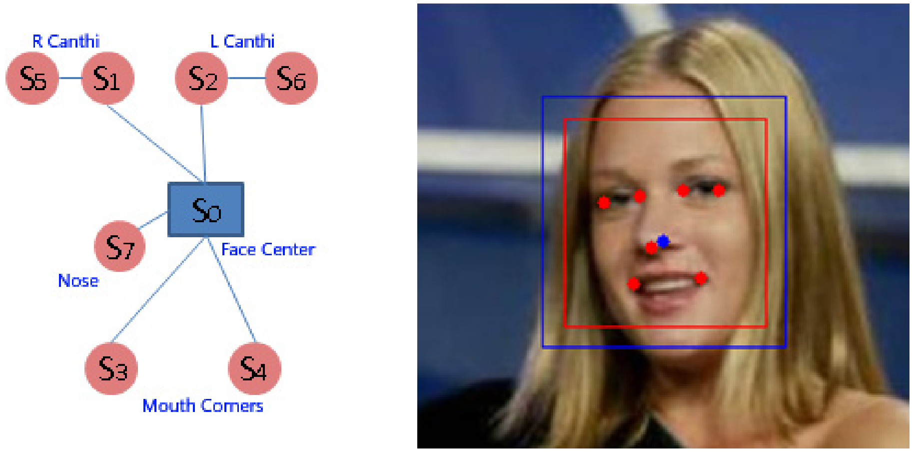

Fig. 2. Left: Underlying Graph of Facial Feature Detected by Flandmark [6], a variation of DPM Right: Results of Flandmark on a sample picture from the LFW Database[7]

2.2 Issues with Side Profiles

Though the DPM approach allows accurate detection of facial features, when used with side profiles, the typical DPM approach is not satisfactory in a number of cases.

One of the first issues faced when using DPM with side profiles is that the feature detectors are usually trained on frontal-like poses. Though individual feature detectors are surprising robust in detection of candidate positions, the features are less stable in its form when the subject is at a wider angler from the ideal frontal profile. Once the features are less stable, the detection process may miss the candidate feature or report the position that may be inaccurate. Features such as the nose and mouth region generally have issues due the form change at wider angles [9].

Another area where the typical DPM approach may not work well is that the underlining graph topology of the facial features and their geometric relationship does not work well with side profiles. In side profiles, there is a possibility that certain features such as one of the eyes may be hidden due to a wide face angle. As in the example when one of the eyes is hidden, the underlying graph is not suitable as a feature is missing and the relationship between should be changed accordingly. In this case, the original approach still attempts to fit the model to the input picture which results in a model over fitting case in which the feature detected are generally inaccurate. The issues discussed with DPM on side profiles are depicted in Figure 3.

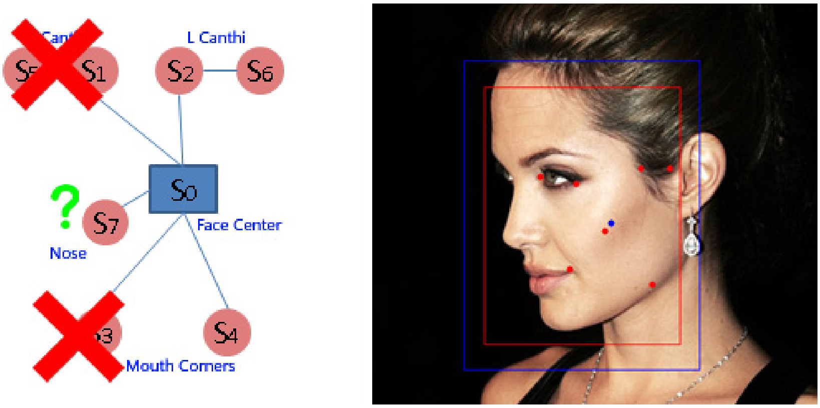

Fig. 3. With a side profile, certain features may be obscured or may change in form. This causes inconsistencies with the default assumptions and causes the model to detect incorrect features due to over-fitting.

3 Proposed System

This section contains details of the proposed system that is implemented. The first subsection contains details of general DPM approaches [4,13] that are selected and utilized in the system, and the second subsection discusses about the variations to DPM approach that are implemented.

3.1 General DPM Approaches

This section describes some of the general DPM approaches that have been selected to use in the proposed system. Assuming picture inputs that are grayscale and of a certain width and height, the system aims to detect the facial features. For the DPM, The configuration of the graph topology is defined by the a graph G = (V,E) where V consists of N features and E is the connection between neighboring features. Each feature is assigned a position s_i in which is the position of the i^{th} feature in the image I. The quality of the feature configuration is then defined by the local feature appearance model based on the match of the feature on position s and the input image, and the deformation cost evaluating the positions related with the neighboring landmarks which is defined by the following equation respectively.

$$F(I,s) = \sum_{i \in V} q_i(I,s_i) + \sum_{(i,j) \in E} g_{ij}(s_i,s_j) \quad (1)$$

The values of q_i and g_{ij} are the combination of of predefined maps and parameter vectors that are learnt from examples which are defined as

$$q_i(I,s_i) = (w_i^q,\ \Psi_i^q(I, si)) \quad (2)$$

$$g_{ij}(s_i,s_j) = (w_{ij}^g,\ \Psi_{ij}^g(si, sj)) \quad (3)$$

The feature descriptor of the local appearance model of the feature Ψ_i^g is computed using the local binary pattern pyramid structure. The local binary pattern pyramid [11] is selected as it provides good performance in texture detection over simpler methods as intensity values and histograms.

For the deformation cost, the quadratic function of the displacement vector is selected [5]. The deformation cost is defined as the following.

$$\Psi_{ij}^g(s_i,s_j) = (dx,dy,dx^2,dy^2),\ (dx,dy) = (x_j,y_j) - (x_i,y_i) \quad (4)$$

3.2 Side Profile Specific Modifications

The first step is to define what set of facial features should be detected in the system. Based on many existing systems, facial features such as the eyes, nose, and mouth are important features to detect. For side profiles, some of the usual features cannot be detected directly. For the proposed system, the selection of the near-eye canthi, nose position, and mouth corner position is selected. The nose position is a vital feature and is important in facial detection. For the eye position, the canthi or the corner of the eyes are important positions. For the canthi, the far canthi are potentially obscured at higher side angles so the near-eye canthi is selected as the feature to be detected. For the mouth corner, the far mouth corner can be difficult to detect when the mouth is open. This is because the far mouth corner is hidden, and the lips form two possible positions for the corner. Based on the features that are selected, it is evident that the typical geometric model and graph topology of a typical DPM approach has to be modified which is displayed in Figure 4.

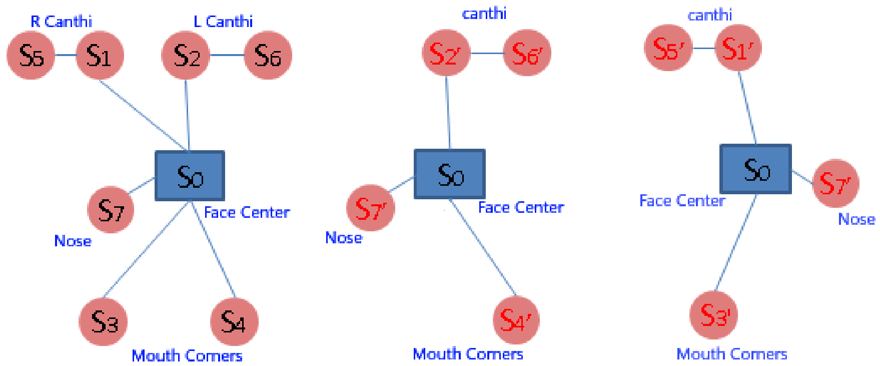

Fig. 4. Left: An example graph topology used in DPM for frontal facial feature detection Middle: A modified topology for side profiles in which subjects are facing to their right Right: A modified/symmetric topology for side profiles in which subjects facing to their left

Another modification is the relaxation of the search space of the feature positions. The typical approach utilizes an AABB bounding box to define a constraint on potential position of the individual features to provide a limited search space and to early prune false positives. The search space is usually constrained to the normal y-axis, and is symmetric in a vertical nature for corresponding feature sets such as the eye canthi. However in side profiles and at certain angles, the features may not be aligned vertically and the y-axis constraints will not be suitable, and a relaxation of the constraint y-axis is permitted to allow a wider range of cascading windows for test in which we allow slight x-axis variation in the search space.

Another modification from the typical approach is the usage of a new face detection approach. Typical approaches utilize Haar's like classifier before enlarging the detected face region for usage. However as commonly available face classifier are trained on frontal images, the approach does not work well with side images. Our proposed system utilizes a HSV color blob model [12] to detect the face region. This works well with a wider range of face region, but has more false positive cases. In our experiment, we select only the positive detections.

Fig. 5. Sample Results from the Proposed System

4 Results and Conclusions

The test cases are built from the complete list of side profile pictures from the Labeled Faces in the Wild (LWF) database [7]. As the number of side profiles in the LFW is small, there has been an additional number of side profile pictures that have been downloaded to increase the test case size. A total of 54 side profiles from the LFW database were selected and an additional 46 side profiles used for the test case. All the selected facial feature of the side profile pictures are marked by human experts.

To compare the results of the proposed system, a popular open source implementation of DPM called Flandmark [6,13] is used for comparison. The proposed system and Flandmark will be used to detect the same facial features that consist of the canthi, nose, and mouth corner. As Flandmark utilizes a frontal haar's classifier to detect the face region which does not detect the side profile well, the face region is manually defined when the detector fails to detect the face region. Also due to difference in features sets, the selection of the closer canthi would be used when comparing the canthi accuracy. To calculate the accuracy, the displacements between the detected and the position have to be calculated. However the pixel Euclidean distance is dependent on the original image size, the normalized coordinate system based on the face region is used instead for both the detected and marked features.

The first calculation done is the feature average mean normalized deviation (F_{AMND}) which is used as an indicator of how accurate each of the feature are which is defined by the following equation

$$F_{AMND} = \frac{1}{M} \sum_{i=0}^{M-1} \| F_i - F'_i \| \quad (5)$$

Where $F = \{F_0, \ldots, F_{M-1}\}$ are the normalized position which is manually marked of each individual feature, $F' = \{F'_0, \ldots, F'_{M-1}\}$ are the positions of each individual feature that are detected by the approach, and M is the total number of test cases.

The second calculation is the average mean normalized deviation (R_{AMND}) which finds the average displacement of all the features in each set. The mean of the deviation of the number of features detected which is defined by F_{no} is calculated first before averaging out the displacement which is defined by the following equation:

$$R_{AMND} = \frac{1}{M} \sum_{j=0}^{M-1} \left(\frac{1}{Fno} \sum_{i=0}^{Fno-1} \| F_{ij} - F'_{ij} \| \right) \quad (6)$$

The last calculation is the average maximum normalized deviation ($R_{MAX\text{-}AND}$) in which selects the feature with the maximum error from each set to figure out what is the potential largest error of each set. $S = \{S_0, \ldots, S_{M-1}\}$ and $S' = \{S'_0, \ldots, S'_{M-1}\}$ contains the set of all the features that are detected, and the maximum deviation is selected from each set before being averaged.

$$R_{MAX\text{-}AND} = \frac{1}{M} \mathrm{MAX}(\| S_i - S'_i \|) \quad (7)$$

The results are compiled, summarized, and are compared side by side in the tables below.

Table 1. Comparing the R_{AMND} and $R_{MAX\text{-}AND}$ between the proposed system and Flandmark

	R_{AMND} [%]	SD R_{AMND} [%]	$R_{MAX\text{-}AND}$ [%]	SD $R_{MAX\text{-}AND}$ [%]
Flandmark	16.6777	10.1151	27.1126	17.2479
Proposed System	3.9029	1.8461	7.3318	4.6764

Table 2. Comparing the F_{AMND} between the proposed system and Flandmark

	Flandmark		Proposed System	
	F_{AMND} [%]	SD F_{AMND} [%]	F_{AMND} [%]	SD R_{AMND} [%]
Canthus1	2.6999	1.4053	12.0171	10.8920
Canthus2	3.8472	2.8434	17.8346	12.5205
Nose	6.3705	5.1268	22.9311	17.4813
Mouth Corner	2.6938	1.9636	13.9280	9.4210

The results of the proposed modified DPM shows improvement in area of the detection of features in side profiles. There is an improvement in the detection of all the facial features when compared with the original system. The improvements are especially evident in the area of the nose position as the features deviates highly from frontal to side profiles. Another cause why the nose position may show a high error rate is potentially from the marked nose position which may fluctuate between test cases which can lead to a higher displacement. One of the interesting results from the system is that the values of Canthus1 and Canthus2, which are the two corners of the eyes, are significantly different. The accuracy of Canthus2 which is the farther eye corner from the camera shows a higher error rate. Though the canthi detection is satisfactory in accuracy, the difference in accuracy points to conclusion that it would be better to train a separate near and far canthus detector in the future work to improve accuracy.

Though the results are satisfactory, the approach focuses detects facial feature on a side profile, and is not a general purpose facial feature detection approach. Currently the authors are working on a general purpose DPM approach that would work seamlessly along a wide range of profiles from frontal to side. Explorations of utilizing techniques such as conditional forest [3] or cascading techniques are currently being explored in which could help solve the issue at hand. Once there is additional progress in the research has been validated, an update to the research is planned.

References

1. Amarapur, B., Patil, N.: The Facial Features Extraction for Face Recognition Based on Geometrical Approach. In: Canadian Conference on Electrical and Computer Engineering, pp. 1936–1939. IEEE (2006)
2. Crandall, D., Felzenszwalb, P., Huttenlocher, D.: Spatial priors for part-based recognition using statistical models. In: CVPR, pp. 10–17 (2005)
3. Dantone, M., Gall, J., Fanelli, G., Gool, L.V.: Real-time facial feature detection using Conditional Regression Forests. In: CVPR (2012)
4. Everingham, M., Sivic, J., Zisserman, A.: Hello! My name is...Buffy" – automatic naming of characters in TV video. In: Proceedings of the British Machine Vision Conference (2006)
5. Felzenszwalb, P.F., Girshick, R.B., McAllester, D., Ramanan, D.: Object detection with discriminatively trained parts based models. IEEE Transactions on Pattern Analysis and Machine Intelligence 99(1) (2009)

6. Flandmark Open-source implementation of facial landmark detector, `http://cmp.felk.cvut.cz/~uricamic/flandmark/`
7. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. University of Massachusetts, Amherst, Technical Report 07-49 (October 2007)
8. OKAO Vision Facial Feature Extraction API, `http://www.omron.com`
9. Setthawong, P., Vannija, V.: Head pose estimation on eyeglasses using line detection and classification approach. In: Papasratorn, B., Lavangnananda, K., Chutimaskul, W., Vanijja, V. (eds.) IAIT 2010. CCIS, vol. 114, pp. 126–136. Springer, Heidelberg (2010)
10. Sivic, J., Everingham, M., Zisserman, A.: Who are you? – learning person specific classifiers from video. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2009)
11. Sonnenburg, S., Franc, V.: COFFIN: A Computational Framework for Linear SVMs. In: Proceedings of ICML (2010)
12. Tathe, S.V., Narote, S.P.: Face detection using color models. World Journal of Science and Technology 2(4) (2012)
13. Uřičář, M., Franc, V., Hlaváč, V.: Facial landmarks detector learned by the structured output SVM. In: Csurka, G., Kraus, M., Laramee, R.S., Richard, P., Braz, J. (eds.) VISIGRAPP 2012. CCIS, vol. 359, pp. 383–398. Springer, Heidelberg (2013)
14. Vatahska, T., Bennewitz, M., Behnke, S.: Feature-based head pose estimation from images. In: 7th IEEE-RAS International Conference on Human Humanoid Robots, pp. 330–335. IEEE (2007)
15. Viola, P., Jones, M.: Fast and Robust Classification using Asymmetric AdaBoost and a Detector Cascade. Neural Information Processing Systems 14, 1311–1318 (2002)
16. Viola, P., Jones, M.: Robust real-time face detection. International Journal of Computer Vision 57(2), 137–154 (2004)