Head Direction Estimation from Silhouette

Amina Bensebaa[1], Slimane Larabi[1], and Neil M. Robertson[2]

[1] University of Sciences and Technology Houari Boumediene,
Computer Science Department, BP 32 El Alia, Algiers, Algeria
slarabi@usthb.dz

[2] Edinburgh Research Partnership in Engineering and Mathematics,
School of Engineering and Physical Sciences, Heriot-Watt University,
Edinburgh, EH14 4AS, UK
n.m.robertson@hw.ac.uk

Abstract. Due to the absence of features that may be extracted from face, heading direction estimation for low resolution images is a difficult task and requires the taking into account all information that may be inferred from human body in image, particularly its silhouette. We propose in this paper a set of geometric features extracted from shape head-shoulders, feet and knees shapes which jointly allow the estimation of body direction. Other features extracted from head-shoulders are proposed for heading direction estimation based on body direction. The constraint of camera position related to proposed features is discussed and results of conducted experiments are presented.

Keywords: Head direction, Body direction, Low resolution, Silhouette.

1 Introduction

Head direction estimation is one of challenging tasks for computer vision researchers especially in case of low resolution images. In case of high and medium resolution images, many approaches has been proposed to solve this problem. A survey may be found in [11]. All of these approaches try to find the most discriminate set of facial features which permit to estimate the pose. The objective to reach for any proposed technique is to verify a set of criteria such as: Accuracy, Monocular, Autonomous, Multi-person, Identity and Lighting invariant, Resolution independent, Full range of head motion and Real time [11].

The problem of estimating heading direction for low-resolution images without adding contextual information requires yet more contributions in order to deal with complex scenes where human are far from the camera. The performance of proposed methods are mainly limited because they are based on extracted features from the head which are very dependent on camera placement and the chosen texture and skin color models depend on the resolution of the head in the image and therefore doesn't work for lower resolution.

In this paper, we will investigate what it can be done from shoulders-head and legs shapes for heading direction estimation in case of low-resolution images. Firstly, a set of features are extracted from shoulders-head and legs shapes

A. Petrosino (Ed.): ICIAP 2013, Part I, LNCS 8156, pp. 340–350, 2013.

and used for inferring body direction. In the next, heading direction is estimated using body direction and features extracted from head-shoulders shape. Section 2 covers the theoretical aspects of body and heading direction estimation based on features extracted from shoulders-head and legs shapes. Experiments are conducted to validate our approach and obtained results are presented in section 3.

2 Prior Works

Face extraction in low-resolution images is an important task in the process of heading direction estimation. Few works have been devoted for this purpose and all present difficulties for detecting faces when the resolution of images decreases [18]. Labeled training examples of head images are used to train various types of classifiers such as support vector machines, neural networks, nearest neighbor and tree based classifiers [13], [3], [4]. The disadvantage of these methods is the requirement of all combinations of lighting conditions and skin/hair colour variations in order to estimate an accurate classification.

Contextual features have been used in addition to visual ones in order to improve the quality of heading direction estimation [9] [8] [1]. Using multiple views camera, Voit et al [17] estimate head pose for low resolution image by appearance-based method. The head size varies around 20×25 and the obtained results are satisfactory due to the use of multiple cameras. Additional contextual information: multiple calibrated cameras and a specific scene allows estimating of absolute coarse head pose for wide-angle overhead cameras by integrating 3D head position [16].

Head-shoulders shape has been studied and many methods have been proposed for the purpose of human detection in images using wavelet decomposition technique and support vector machine [14] or background subtraction algorithm [12]. In other side, head-shoulders shape has been used for human tracking and head pose estimation. In [12], the direction of head movements is detected and tracked throughout video frames. Templates are captured for a specific position of the camera (mounted sufficiently high above to provide a top-view of the scene) without using all head poses.

Legs shape may contribute for heading-direction estimation. Indeed, detectors on the lower parts of the body has been introduced in many works for human body pose calculation and human action recognition [15]. Legs shape has been also used for human segmentation where body parts particularly the legs are modeled in order to detect and segment human [10]. The proposed approach is based on the matching of part-template tree images hierarchically proposed and used initially in [6], [7].

Our main contributions in this work are:

- The processing of images of far-field video when there is a difficulty to locate some features used such as skin, face.

- A set of geometric features are proposed that allow the estimating of body's and head's directions,
- Such features may be extracted even for far-field video because they concern only the silhouette.
- Possibility to extract these features whatever the position of the camera except for the top view.

3 Basic Principle of the Method

Assuming that silhouettes of humans are extracted from images of low resolution, our aim is to estimate the directions of body and head. Geometric features are extracted from silhouette due to the absence of other features that may be extracted from head (face and hair) for such images. We will focus in this paper on shapes of head, shoulders, knees and feet which may be considered as a good features to achieve this task. Body direction is firstly estimated using features extracted from head-shoulders, knees and feet shapes. Secondly, knowing body direction, head direction is inferred from features of head-shoulders shape.

3.1 Geometric Features from Silhouette

A shape leg is a part of silhouette which plays a dominant role for body direction estimation from image. Indeed, our visual system is able to infer body direction seeing only the outline shape legs (and/or) head-shoulders shape (see figure 1). We propose three determinant cues of shapes legs and head-shoulders that allow inferring body direction when they are extracted from outline shape. These features can't be computed for a fixed top down camera giving blobs of persons where head-shoulders can't be extracted from the rest of the silhouette.

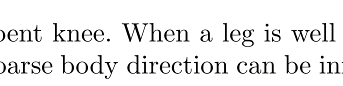

Fig. 1. Some shapes of legs for which it is easy to infer body direction

The **first one** is the bent knee. When a leg is well separated from the other and the knee is bent, a coarse body direction can be inferred without ambiguity. Figure 2.a illustrates an example of shape legs where feet are hidden. Our visual system can easily give an estimate of body direction because the feet have limited possibilities of poses. Figure 2.b illustrates the correct poses and the directions can be inferred using the feet shapes, however figure 2.c shows impossible situation. The directions of the lines joining inflexion points of the same leg are used to infer the body direction.

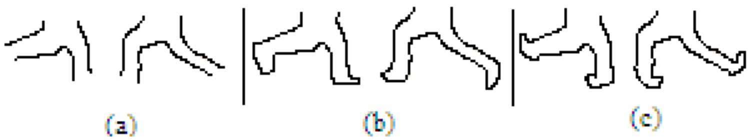

Fig. 2. Shapes of legs with inflected knee

The **second one** is the direction of shape foot. Indeed, our visual system encounters difficulties by looking at legs shapes without feet and can't estimate body direction for many configurations even if the body is moving and legs are well separated but without inflexion of knees. For example, seeing to the outlines of figure 3.a, without feet we can't recognize to what direction body is moving. This ambiguity is clear seeing at the original shapes (see figure 3.b) and at new shapes obtained drawing feet (see figure 3.c). The base lines of the feet are good features because they indicate the body direction. Their use is explained in the following subsection.

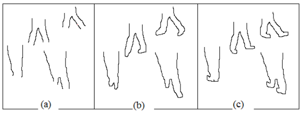

Fig. 3. Ambiguity in body direction estimation in case of missed shape feet

The **third feature** concerns the variation of silhouette's width along the shape head-shoulders. The ratio of the width of the upper part (head) and the lower part (shoulders) is related to the angle of rotation. We noticed that there's an opposite relationship between the ratio and the orientation angle.

3.2 Inferring Body Direction

Body Direction Estimation Using Feet's Features. This task consists to split the lower human shape into separated legs, separated lower legs or grouped legs (The two first cases include the case where the knee of one leg may be bent). This is done by the concavity points located on the outline in the lower part of the outline indicating the feet.

We associate to each foot a **base line** defined by two extremities of the foot located between the heel and the toes. The outline of lower part is processed in order to determine the baseline of the feet located between the heel and the toes. Firstly, high convexities points Cv_1 and Cv_2 characterizing the outline foot are located (see Figure 4(a)). Secondly, the last point of interest Cc representing a high concavity on this outline is located, such as the distances $CcCv_2$ is minimal. The convex point that represents toes, will be the closest point to the concave point of the feet outline, the other convex point will obviously correspond to the heel. Thus the base line joins the two convexities of the foot and the orientation of feet corresponds to the vector carried by the feet base line.

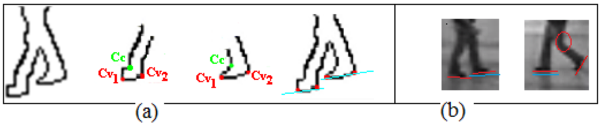

Fig. 4. (a) Steps of body direction estimation based on foot directions, (b) Body orientation from feet (In red color the feet orientations of foot and in blue the body orientation)

Applying the 2D quasi-invariant, the angle between the two vectors measured in 3D-space varies slowly in the image as viewpoint varies [2]. As in the scene the disposition of foot vectors is restricted by the human physic constraints, it will be the same case in image plane; the body direction is inferred as the average of foot directions. Once the base lines of feet are extracted, body orientation is computed as the resultant vector of the two orientations. When one foot isn't put on the ground, which correspond to a high curvature of the knee, the resultant vector will have the direction of the base line of the other foot (see figure 4(b)).

Body Direction Estimation Using Knee's Features. Extraction of curvature points consists to find the best concave or convex pixels of the lower part of the silhouette. Using the Chetverikov's algorithm [5], p is selected as curvature point if the angle defined by the three pixels: p^-, p, p^+ is lower that a given threshold α. Pixels p^-, p^+ are located pixels before and after the considered pixel p at d pixels (see figure 5). Default values of d, α are $7pixels, 150$.

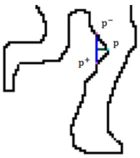

Fig. 5. Location of curvature points on outline legs

Many types of knees inflexion may be located (see figure 6) and body direction is considered as the direction defined by the concave point to the convex one of the bent knee. Only the direction left towards right and inversely will be considered.

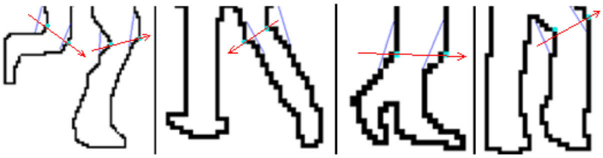

Fig. 6. Some cases of knee inflexion and the inferred direction of the knees

Body Direction Estimation Using Head-Shoulders Features. Body direction can inferred using the head-shoulders geometry. Indeed, when human body is rotating, the ratio R_w of the widths of head and shoulders are estimated as follow:

For the rotation angle in the intervals : $[0°, 15°]$, $[15°, 30°]$, $[30°, 45°]$, $[45°, 60°]$, $[60°, 75°]$, $[75°, 90°]$, the ratio R_w belongs respectively to the intervals: $[1.82, 4]$, $[1.70, 1.81]$, $[1.61, 1.69]$, $[1.51, 1.60]$, $[1.36, 1.5]$, $[1.4, 1.5]$.

Applying the algorithm of D. Chetverikov [5], the two concave points (left and right) delineating the head and the two convex points (left and right) extremities of shoulders are located. Head is separated by locating the pixels having the minimum angle among the selected candidate points. The second extremities of shoulders are located as convex pixels with high curvature along of the outline head-shoulders (see figure 7).

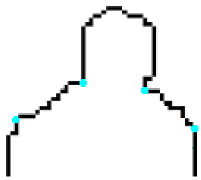

Fig. 7. The second extremities of shoulders

3.3 Inferring Head Direction from Head-Shoulders Shape

We assume that body direction is estimated based on the three features proposed above. In order to estimate head direction, two features are extracted from head-shoulders outline.

Features Extraction. The first feature concerns the lengths of shoulders S_L, S_R on shape head-shoulders. When the end of the neck isn't visible on one side due to shoulder's occlusion by head, the beginning of the shoulder is considered as the point of high curvature on head-shoulders outline.

The lengths of shoulders are important cues for both head and body directions estimation and the difference between lengths of S_L, S_R arises from one of the following configurations:

- Depending on the camera and body positions, the head can occludes a part of one shoulder and then decreases the shoulder length. For example, when the camera is on top at the right or at the left of the person (see figure 8).
- When human body is rotating, one of shoulders becomes less visible. This occurs for example when the camera is on top even if the person is in front of the camera. In this case, length of one shoulder decreases until that the two sides of the shape head-shoulders don't correspond to shoulders.

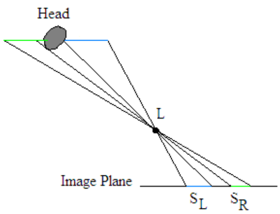

Fig. 8. Cases of occlusion of shoulder by head

Consequently, when body direction and head are in front to the camera, the lengths $L(S_L), L(S_R)$ of shoulders are identical. Otherwise, when the head is rotating or when body is at the lateral side of the camera, this equality isn't verified because in both cases the head occludes a part of one shoulder (see figure 8). We proved geometrically that without occlusion by head, the lengths of one shoulder decreases when body is rotating.

The second feature which completes the first one, concerns the occluded parts of shoulders that permit to estimate head rotation. Let I be the intersection point of the lines joining extremities of shoulders S_L and S_R (see figure 9). When body and head are in front to the camera, the distances d_L, d_R from I to shoulders are identical in the scene and in image. However, when head or body are rotating, these distances are different in image because a part of shoulder is occluded by head and thus in image the distance d_L or d_R includes the occluded segment of the shoulder and a part of the neck. The distances d_L, d_R will be used to infer the head direction.

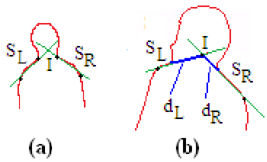

Fig. 9. Intersection of shoulders in case where (a) body and head are in front, (b) body and head rotating

Coarse Estimation of Head Direction. Head direction is estimated assuming body orientation, the difference ΔL between the images of (S_L), (S_R) and the difference Δd between the images of d_L, d_R are computed. When body is in the center of the field of view or at the left, table 1 gives the results obtained of head direction applying a geometric reasoning depending on the values of ΔL, δd and body direction. For example, figure 10 illustrates the variation of ΔL, δd in case where human in the center of the field of view of the camera. The case where body is at the right is symmetrical to the case where body is at the left.

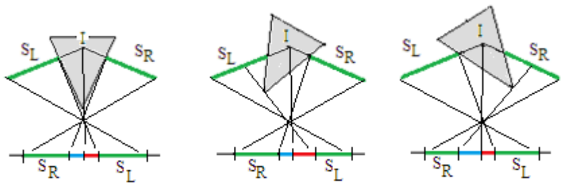

Fig. 10. Different poses of head where images of d_R, d_L are illustrated with blue and red color in case where human is in the center of the field of view

Table 1. Head direction inferred in cases where body is in front and at the left

Body in the center	$\delta d = 0$	$\delta d > 0$	$\delta d < 0$
$\Delta L = 0$	Head in Front	Not possible	Not possible
$\Delta L < 0$	Not possible	Rotation to left	Not possible
$\Delta L > 0$	Not possible	Not possible	rotation to right
Body at the left	$\delta d = 0$	$\delta d > 0$	$\delta d < 0$
$\Delta L = 0$	Low Rotation to right	Not possible	Not possible
$\Delta L < 0$	Not possible	Head in front or rotating to left	Not possible
$\Delta L > 0$	Not possible	Not possible	Hight Rotation to right

3.4 Study of the Camera Position Constraint

As we are interested in this work to images of low resolution which means a far field of view, the camera may be:

- Fixed at the top and far from the scene. In this case, none from the features: head, shoulders, legs and feet can't be located using the blob representing human.

- Fixed so as its optical axis is oblique or horizontal towards the scene. In this case, whatever the position of the camera relatively to human in the scene: in front or at the lateral position, its head-shoulders, legs and feet are viewed. Consequently, the availability of the proposed features depends only on the pose, which means that inflexion of knees or feet base lines may be missed, however the presence of the head-shoulders outline is required.

4 Results

We applied our method on some real images. Firstly silhouettes are extracted and body direction is firstly computed. In the next, head direction is estimated. We used all features extracted from head-shoulders, feet and knees outlines.

Figure 11 illustrates some poses, extracted silhouettes and computed body directions. Body direction is computed using the ratio R_w equal to $2.6, 2.89, 2.25, 1.33, 1.36, 2.27, 2.09$ giving the intervals of head direction equal to: $[0\,°, 15\,°], [0\,°, 15\,°], [15\,°, 30\,°], [75\,°, 90\,°], [75\,°, 90\,°], [15\,°, 30\,°], [0\,°, 15\,°]$. We can see that the computed body direction for the two last poses $(f), (g)$ are done using only the first feature which can't differentiate if the body in in front or of back with regard to the camera. The orientation of feet, when are located in the image, eliminate the ambiguity (in front or of back).

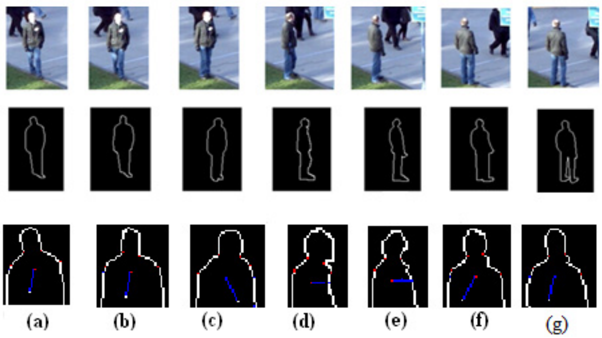

Fig. 11. Some poses and extracted silhouettes and the computed body directions based on R_w values

The combination of features used for body direction depends on what can be extracted. The features extracted from feet and knees are more strong than those extracted from head-shoulders. Figure 12 illustrates the results obtained when inflexion of knees are used in addition of the ratio R_w.

Head direction estimation is based on estimated body direction and the values of d_L, d_R computed using head-shoulders outline. Figure 13 summarizes this combination of features and shows that a good estimation is made even if the images are of low resolution.

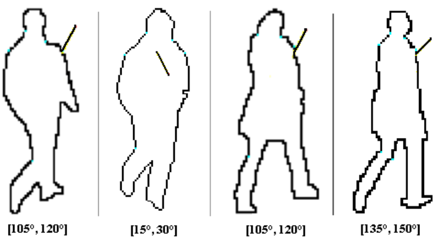

Fig. 12. Body Orientation using knee inflexion and R_w ratio

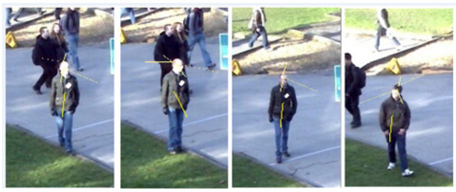

Fig. 13. Head and body directions from combined features

5 Conclusion

We proposed in this paper a method for head direction estimation based on geometric features extracted even if images are of low resolution. Body direction is inferred from features extracted from outlines of knees and feet and head-shoulders. This direction is used in addition to features extracted from outlines of head-shoulders for estimating head direction. The proposed method has been applied on real images and achieves good estimation of head direction. Also, the features extracted are independent from camera pose, except the top view where head-shoulders, knees and feet can't be located on human shape.

References

1. Ba, S.O., Odobez, J.M.: Multiperson Visual Focus of Attention from Head Pose and Meeting Contextual Cues. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(1), 101–116 (2011)
2. Binford, T.O., Levitt, T.S.: Quasi-invariants: Theory and exploitation. In: Proceedings of DARPA Image Understanding Workshop, pp. 819–829 (1993)
3. Benfold, B., Reid, I.D.: Colour Invariant Head Pose Classification in Low Resolution Video. In: Proceedings of the 19th British Machine Vision Conference (2008)
4. Benfold, B., Reid, I.D.: Unsupervised Learning of a Scene-Specific Coarse Gaze Estimator. In: Proceedings of the International Conference on Computer Vision (ICCV), pp. 2344–2351 (2011)
5. Chetverikov, D.: A Simple and Efficient Algorithm for Detection of High Curvature Points in Planar Curves. In: Petkov, N., Westenberg, M.A. (eds.) CAIP 2003. LNCS, vol. 2756, pp. 746–753. Springer, Heidelberg (2003)

6. Gavrila, D.M.: The Visual Analysis of Human Movement: A Survey. Computer Vision and Image Understanding 73(1), 82–98 (1999)
7. Gavrila, D.M.: A Bayesian, Exemplar-Based Approach to Hierarchical Shape Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(8), 1408–1421 (2007)
8. Lanz, O., Brunelli, R.: Joint Bayesian Tracking of Head Location and Pose from Low-Resolution Video. In: Stiefelhagen, R., Bowers, R., Fiscus, J.G. (eds.) RT 2007 and CLEAR 2007. LNCS, vol. 4625, pp. 287–296. Springer, Heidelberg (2008)
9. Launila, A., Sullivan, J.: Contextual Features for Head Pose Estimation in Football Games. In: International Conference on Pattern Recognition (ICPR 2010), Turkey, pp. 340–343 (2010)
10. Lin, Z., Davis, L.S.: Shape-Based Human Detection and Segmentation via Hierarchical Part-Template Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(4), 604–618 (2010)
11. Murphy-Chutorian, E., Trivedi, M.M.: Head Pose Estimation in Computer Vision: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(4), 607–626 (2009)
12. Ozturk, O., Yamasaki, T., Aizawa, K.: Tracking of humans and estimation of body/head orientation from top-view single camera for visual focus of attention analysis. In: IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), pp. 1020–1027 (2009)
13. Robertson, N., Reid, I.D.: Estimating Gaze Direction from Low-Resolution Faces in Video. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3952, pp. 402–415. Springer, Heidelberg (2006)
14. Sun, Y., Wang, Y., He, Y., Hua, Y.: Head-and-shoulder detection in varying pose. In: Wang, L., Chen, K., S. Ong, Y. (eds.) ICNC 2005. LNCS, vol. 3611, pp. 12–20. Springer, Heidelberg (2005)
15. Singh, V.K., Nevatia, R., Huang, C.: Efficient Inference with Multiple Heterogeneous Part Detectors for Human Pose Estimation. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part III. LNCS, vol. 6313, pp. 314–327. Springer, Heidelberg (2010)
16. Tian, Y.L., Brown, L., Connell, C., Sharat, P., Arun, H., Senior, A., Bolle, R.: Absolute head pose estimation from overhead wide-angle cameras. In: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, AMFG 2003, pp. 92–99 (2003)
17. Voit, M., Nickel, K., Stiefelhagen, R.: A Bayesian Approach for Multi-view Head Pose Estimation. In: IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, pp. 31–34 (2006)
18. Zheng, J., Ramírez, G.A., Fuentes, O.: Face detection in low-resolution color images. In: Campilho, A., Kamel, M. (eds.) ICIAR 2010. LNCS, vol. 6111, pp. 454–463. Springer, Heidelberg (2010)