The Imaging Science Journal

ISSN: 1368-2199 (Print) 1743-131X (Online) Journal homepage: http://www.tandfonline.com/loi/yims20

Fast and accurate direction estimation of moving pedestrians

Amina Bensebaa & Slimane Larabi

To cite this article: Amina Bensebaa & Slimane Larabi (2018): Fast and accurate direction estimation of moving pedestrians, The Imaging Science Journal

To link to this article: https://doi.org/10.1080/13682199.2018.1542781

Published online: 15 Nov 2018.

THE IMAGING SCIENCE JOURNAL
https://doi.org/10.1080/13682199.2018.1542781

THE RPS ROYAL PHOTOGRAPHIC SOCIETY
Taylor & Francis Taylor & Francis Group

RESEARCH ARTICLE

Fast and accurate direction estimation of moving pedestrians

Amina Bensebaa and Slimane Larabi

Computer Science Department, University USTHB, Algiers, Algeria

ABSTRACT

This work proposes a simple yet efficient way to estimate pedestrians flow direction based on videos from still cameras. It does that by localizing the extremities of head and feet of silhouettes and fitting them to lines. As the previous in three-dimensional space of these lines are parallel, their intersection point is the vanishing point. Using the computed vanishing point and two internal camera parameters, the horizontal direction of moving pedestrian is determined. Our method competes for the state-of-the-art methods and achieves a high rate accuracy for direction classification.

ARTICLE HISTORY
Received 16 December 2017
Accepted 26 October 2018

KEYWORDS
Projective geometry; vanishing point; direction; video sequence

Introduction

Locating pedestrians on video sequence frames and estimating their directions is a crucial step for computer vision applications. Many contributions have been proposed and the challenge still remains due to external conditions that degrade the images quality especially for low-resolution images.

In this paper, we address the problem of estimating the direction of moving single pedestrians. The scene is depicted by a series of images taken by a single camera. Movements of head and feet of pedestrian's silhouette serve for determining the direction of movement of that person. Each movement (top and bottom) is described by a line. In order to be robust against the small individual movements of the head and the feet, the positions of the head and feet, respectively, are fitted to a top and a bottom line. According to projective geometry, these lines, that are parallel in 3D, intersect in a vanishing point in the image plane. Using the estimated vanishing point, the principal point coordinates and the focal length, the direction of the movement of the person is determined. Figure 1 illustrates this process applied to a video sequence.

Contributions

Assuming that frame segmentation may be poor (shadow not removed, additional or missing parts of silhouettes), our aim in this paper is to estimate the direction of pedestrians in such conditions and to improve the results of the state-of-the-art for single moving pedestrian. Our contributions are multiple:

- Ability to estimate the direction of pedestrians even for low-resolution images.
- Ability to apply the method to estimate the direction of rigid objects.
- Ability to take into account the non-planarity of the ground and walking on the stair.
- Fast and accurate direction estimation surpassing the results of the state-of-the-art.

Context of applications

The proposed method locates pedestrians crossing or walking and estimates their directions, is intended to:

- Accident avoidance system for smart cars (see Figure 2).
- Pedestrian monitoring in traffic control systems.
- Prediction for tracking and motion analysis in visual surveillance.
- For video surveillance, knowing the direction of a person is crucial to prevent access to monitored areas. Also, when abnormal event occurs, providing the direction of escaped person is a major information in order to track it by cameras or police cars.
- Using the reference points (top and bottom) may be extended and used for direction estimation of rigid objects such as vehicles. When a car moves, the top and bottom points define in 3D parallel lines. Their projections in images define the vanishing point associated to the direction of that car in 3D. For intelligent systems embedded in vehicles, directions of neighbouring vehicles are essential for navigation (see Figure 3).

Paper roadmap

In the next section, we present essentially the recent published methods related to our work. We explain the basic fundamentals of our method and the

CONTACT Slimane Larabi slarabi@usthb.dz; slimane.larabi@gmail.com

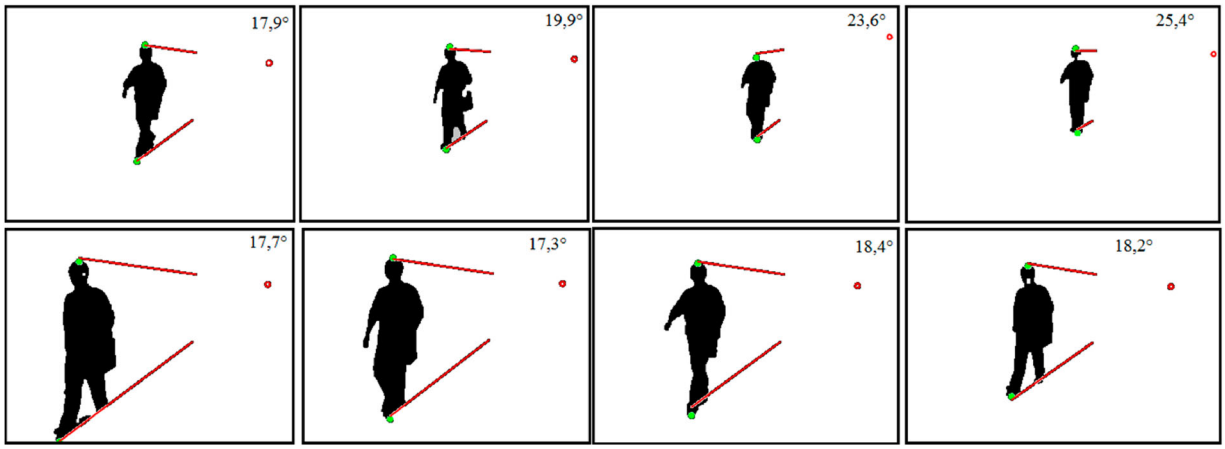

Figure 1. Direction estimation of pedestrians: The basic principle proposed in this work is to locate the top and bottom points of silhouette's pedestrian along the video sequence and to fit them by regression. The intersection of the two lines provides a vanishing point (circle) associated to pedestrian's direction. Note that the ground truth direction is 18° and the accuracy of computed direction (value on figure) increases during walk time.

derived algorithm in Section 3. In Section 4, we explain how the proposed algorithm is adapted to face to complex scenes (non-planar ground, stair) and multiple trajectories. We present the conducted experiments on available datasets and we discuss the obtained results in Section 5. A discussion related to method's efficiency is given in Section 6. We conclude this paper in Section 7.

Related works

As a rough direction of the body is sufficient for many computer vision applications, some of the proposed methods are based on classification techniques.

For 16 (respectively 8) directions, a score of 90% (respectively 64%) of correct recognition is obtained using the Support Vector Machine (SVM) in [4] (respectively. Adaboost classifier [5]). HoG feature extracted from each people and used as an input to an array of binary classifiers trained on a set of discrete orientations gives a of score 70% for correct classification of 8 directions [6].

An appearance-based classifier [7] estimates the pedestrian's orientation. Reliable motion direction is determined acting as pre-estimated person orientation to update the appearance-based classifier.

Others classification techniques have been used such as Random Decision Forests [8] which yields results comparable to those of state-of-the-art and baseline methods. Hidden Markov Model is used for classification, 90% of accuracy is achieved on CASIA-B dataset [9] by Raman et al. [10]. Deep-learning approach proposed by Raza et al. [11] outperforms existing classifiers and achieves accuracy 0.92 for full-body orientation estimation using only eight orientation classes.

The vanishing points have been used in [12] for camera calibration, and for pedestrian direction estimation [13,14] assuming that the line joining the head positions of a pedestrian at two-time instances is parallel to the line joining the feet (or bottom) positions at these same corresponding time instances. Similarly, the line joining the head to the feet in a frame is parallel to the line joining the head to foot the next frame.

The vanishing point is an interesting idea because it is associated to the horizontal direction of walking pedestrian. Based on these reference points, we investigated in this work what it can be done from extracted noisy silhouettes.

Proposed approach

Our method begins by background estimation and silhouette's pedestrian extraction taking into account the shadow elimination. The next step is locating the top

Figure 2. For smart cars, directions of pedestrians are required for accident avoidance system. Source: Left [1], Right [2].

Figure 3. Direction of moving cars can be inferred from silhouettes from the video sequence. Source [3].

and bottom extremities of the silhouette and their fitting into lines. We demonstrate that the horizontal pedestrian direction is inferred from the vanishing point in the image plane as the intersection of the fitted lines.

The corresponding algorithm is then presented and adapted to take into account the non-planarity of the ground and the pedestrian's motion on the stair.

Silhouette extraction

Background is estimated periodically in order to tackle weather change for outdoor scenes and lighting change for indoor scenes. To do this, the luminance of each pixel is evaluated as the average of its values in a set of frames. Foreground is extracted using a chosen threshold. Applying the method proposed in [15] the shadow is removed. Figure 4 illustrates an example of shadow elimination on PETS dataset [16].

Walking Direction of Single Pedestrians from 3D to 2D

Let $\omega(u_\omega, v_\omega)$ be the vanishing point located on the image plane associated to the direction of $((\Delta^h), (\Delta^f))$ defined by pedestrian bodies parts (**h**ead and **f**eet) along the motion trajectory (see Figure 5).

We will consider the pedestrian's direction as the value of the angle α relatively to the axis $\overrightarrow{PY}$, where $(PXYZ)$ is the associated camera reference.

The value of α is computed using ω, the focal length f and the u-coordinates of the principal point $P(u_O,\ v_O)$ as follow:

$$\alpha = \tan^{-1} \frac{u_\omega - u_O}{f} \tag{1}$$

Determining the values of u_O and f

The values of u_O, f needed for calculating the α angle, may be available either:

- Camera calibration,
- Parameters evaluation by performing a leaning stage using at least two pairs of parallel lines having two known different directions α_1, α_2 giving two vanishing points ω_1, ω_2 in the image plane. The values of u_O, f are computed as follow.

$$f \times \tan \alpha_1 = (u_{\omega_1} - u_O) \tag{2}$$

$$f \times \tan \alpha_2 = (u_{\omega_2} - u_O) \tag{3}$$

If $\alpha_1 \neq 0$ and $\alpha_2 \neq 0$, the values of u_O, f are estimated by solving the system of linear Equations (2) and (3). Otherwise, if $\alpha_1 = 0$, then $u_{\omega_1} = u_O$ and it is necessary to have $\alpha_2 \neq 0$ required to compute f.

Finding pair of parallel 3D lines in the scene

Images of walking pedestrian contain geometric features useful for inferring information about 3D scene, especially features related to projective geometry assuming that image may be considered as a perspective projection of the 3D scene.

If we look at Figure 6, we can see that when pedestrian moves from one position to another on the

Figure 4. Current frame from PETS2006 [13] and foreground detection, shadow is removed. Source [16].

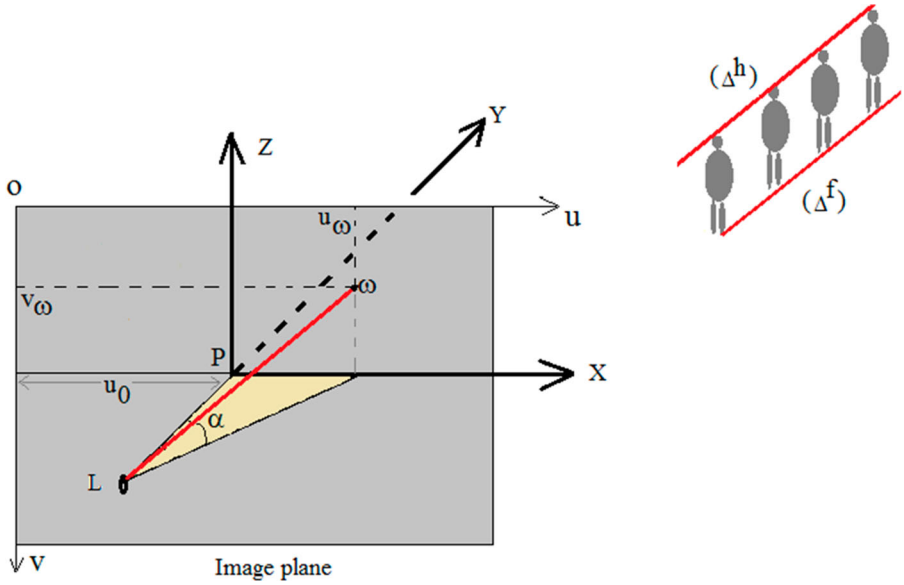

Figure 5. Vanishing points ω associated to the direction of $((\Delta^h), (\Delta^f)$.

same ground, the two lines defined by the limits of the heads and the limits of feet may be considered as parallel to the ground [13,17]. In the image plane, their projections (δ^h, δ^f) intersect at the vanishing point ω.

In theoretical case, joining the bottom points of pedestrian's images will define a line if they correspond to the same 3D point of pedestrian's feet. This is an unrealistic situation, because feet in 3D is moving and then the bottom points are unmatched points (see Figure 6).

In the same case, the head pose may change. Consequently, in image the two lines will be determined by minimizing the distance from the extremities (top, bottom) to the considered line (the best-fitting linear regression model). Note that the more numerous are the used silhouettes in this process, the more accurate is the fitted line in the image.

The Naive Algorithm
Begin
F_i, $i = 0..k$ are the frames to be processed.
Focal length (f) and u-coordinate of the principal point u_O are computed previously.
- $i = 0$ # the order number of the first frame F_0.
- Locate the top and bottom points p_0^h, p_0^f of the pedestrian's silhouette for the frame F_0

For each F_i ($i = 1..k$)
Do
1. Locate the top point p_i^h and bottom point p_i^f of the pedestrian's silhouette on F_i.
2. Fit via linear regression the top points p_j^h ($j = 0..i$) to new line (δ^h)
3. Fit via linear regression the bottom points p_j^f ($j = 0..i$) to new line (δ^f)
4. Compute the vanishing point $\{\omega\} = \delta^h \cap \delta^f$.
5. Compute the value of $\alpha = \tan^{-1}\left(\frac{u_\omega - u_O}{f}\right)$

# α is the estimated direction for the frame F_i
EndFor
End.

Direction computation from images sequence: Algorithm

The first step is the computation of camera parameters u_O, f as explained in subsection 3.2. In order to get accurate values of u_O, f, we will take more than two directions. More details are given in the validation section. Once these parameters are estimated, we compute the direction for each frame using the Equation (1).

Adapting the method for complex scenes

Adapting the method to occlusion, segmentation errors, non-planar ground, direction change

It is clear that if we use all frames to determine the lines fitting the top and bottom points, the intersection point (vanishing point) is sensible to occlusion, segmentation error, direction change and non-planar ground. We can see in Figure 7(a) that head of

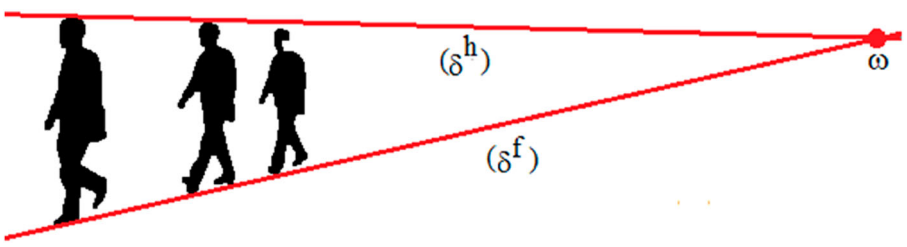

Figure 6. Images of the parallel lines in 3D giving the vanishing point in 2D.

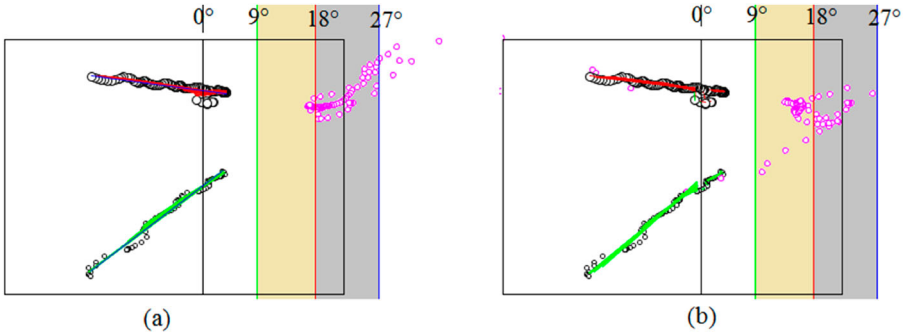

Figure 7. Vanishing points computed applying the naive algorithm (a) and improved algorithm (b).

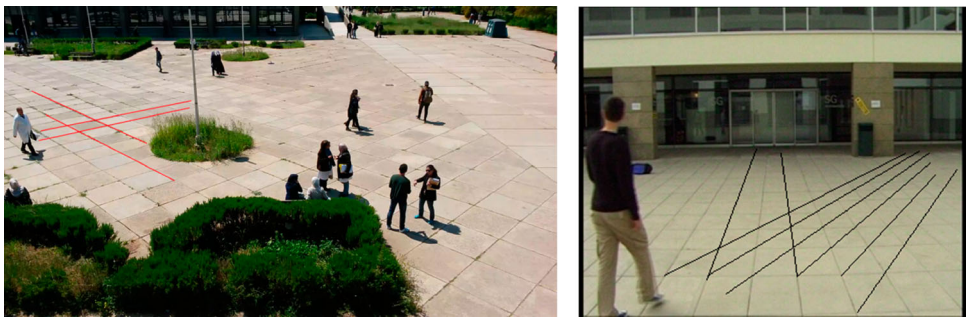

Figure 8. Parallel lines located on the ground for MGP01 and Campus datasets. Source: Left [17], Right [http://perso.usthb.dz/~slarabi/MGP01.html].

pedestrians are missing from frames 46–55 due to occlusion, the top points, in this case, are taken from shoulders. The previous fitted line for top points becomes far from the new top points.

However, instead to consider for the current frame all previous frames, we consider only the previous ones such that the slope is approximately stable for the new fitted line. We verify if the located top point is near from the fitted line for the previous frames, then the current frame will be included in the fitting of the new line. Otherwise, the current frame will be considered as first one of a new fitting process. This will allows to get good estimation in case of missing parts of silhouette, direction change and non-planar ground, because the top and bottom points will define two parallel lines. Figure 7(b) illustrates how the fitted lines are corrected giving better results.

Table 1. The mean error μ and standard deviation σ for Casia-A for the delay 50, 30, 10 frames.

	0	45°	90°
$\mu(50)$	4.73	6.06	1.9
$\sigma(50)$	3.14	10.14	5.81
$\mu(30)$	5.73	6.20	2.22
$\sigma(30)$	5.07	10.20	15.5
$\mu(10)$	8.87	6.86	3.74
$\sigma(10)$	11.19	10.88	16.16

Table 2. The mean error μ and standard deviation σ for Casia-B for the delay 50, 30, 10 frames.

	0°	18°	36°	54°	72°	90°
$\mu(50)$	7.73	1.71	1.57	4.11	1.28	1.88
$\sigma(50)$	0.67	2.69	1.69	3.05	1.57	1.18
$\mu(30)$	8.27	2.29	2.27	6.86	1.63	3.71
$\sigma(30)$	1.49	3.21	3.10	6.42	2.06	2.25
$\mu(10)$	11.87	4.29	4.55	8.73	3.75	5.75
$\sigma(10)$	11.48	8.41	8.95	7.85	5.29	5.4

The new Algorithm and Complexity

The New Algorithm
Begin
F_i, $i = 0..k$ are the frames to be processed.
Focal length (f) and u-coordinate of the principal point u_O are computed previously.
- $i_{start} = 0$ # the order number of the starting frame.
- start_again:
{
- $i = i_{start}$
- Locate the top points p_i^h, p_{i+1}^h and the bottom points p_i^f, p_{i+1}^f of the pedestrian's silhouette for the frame F_i, F_{i+1}
- Fit via linear regression the top points to the line δ^h
- Fit via linear regression the bottom points to the line δ^f
- Compute the vanishing point $\{\omega\} = \delta^h \cap \delta^f$.
- Compute the value of $\alpha = \tan^{-1}\left(\frac{u_\omega - u_O}{f}\right)$
α is the estimated direction for the frame F_i, F_{i+1}
- i=i+2
}
For each F_j ($j = i..k$)
Do
1. Locate the top point p_j^h and bottom point p_j^f of the pedestrian's silhouette on F_j.
2. If(p_j^h is near from (δ^h))
Then Fit via linear regression the top points p_k^h ($k = i_{start}..j$) to new line (δ^h)
Else $i_{start} = j$, Goto start_again
EndIf
3. If(p_j^f is near from (δ^f))
Then Fit via linear regression the bottom points p_k^f ($k = i_{start}..j$) to new line (δ^f)
Else $i_{start} = j$, Goto start_again
EndIf
4. Compute the vanishing point $\{\omega\} = \delta^h \cap \delta^f$.
5. Compute the value of $\alpha = \tan^{-1}\left(\frac{u_\omega - u_O}{f}\right)$
# α is the estimated direction for the frame F_j
EndFor
End.

Table 3. Confusion matrix for Casia-B dataset with delay 0, 20, 40 frames. Values given are percentages.

Directions (delay)	D1	D2	D3	D4	D5	Accuracy	Accuracy [7]
D1(0 f)	74.78	5.54	6.22	3.13	10.3	74.78	93.05
D1(20 f)	83.97	2.03	1.26	1.86	10.86	83.97	
D1(40 f)	93.34	0.48	0	0.34	5.82	93.34	
D2(0 f)	0.87	91.99	5.83	0.77	0.51	91.99	90.99
D2(20 f)	0.19	98.09	1.64	0.06	0	98.09	
D2(40 f)	0	100	0	0	0	100	
D3(0 f)	0	4.17	95.44	0.22	0.16	95.44	87.97
D3(20 f)	0	1.08	98.73	0	0	98.73	
D3(40 f)	0	0	100	0	0	100	
D4(0 f)	0	0.06	23.38	76.35	0.19	76.35	89.75
D4(20 f)	0	0	14.64	85.35	0	85.35	
D4(40 f)	0	0	5.66	94.33	0	94.33	
D5(0 f)	1.85	1.53	2.23	14.65	79.71	79.71	92.60
D5(20 f)	0	0.24	0.24	7.31	92.17	92.17	
D5(40 f)	0	0	0	1.89	98.1	98.1	
Our(0 f)			Balanced Accuracy			83.65	90.87
Our(20 f)						91.66	
Our(40 f)						97.15	

The complexity of the algorithm is $O(mk)$, it depends on m (the number of pixels of outline silhouette) and on k (the number of frames). If we consider that $m \sim n$, where $n \times n$ defines the image size, then the complexity is $O(n)$.

Validation

Values of focal length and principal point coordinate

Four datasets CASIA-A, CASIA-B [9], MGP01 [18] and EPFL Campus [19,20] have been used. For each one, the focal length and principal point coordinates f, u_O are estimated.

For CASIA-A and CASIA-B datasets, from the available ground truth, we used video sequences of two known directions and the values (325, 187) are obtained for f, u_O. For MGP01 and Campus datasets, we located on images two pairs of lines such that their previous are parallel in 3D as illustrated by Figure 8. The computed parameters are respectively (2141, 964) and (206, 181).

Accuracy according to the delay: Results

Applying our method, the estimated direction depends strongly on the number of pairs of silhouette's extremities (top, bottom) because the more numerous are the points; the more accurate are the fitted lines.

We studied using Casia-A and Casia-B datasets the accuracy of estimated direction when is done after a certain delay. The obtained results (see Tables 1 and 2) indicate that after 50 frames (3 sec of walking) are sufficient to get an accurate estimation.

Results for Casia-A and Casia-B datasets

For these datasets, we consider that if the error of estimated direction is less than 9°, the direction is correctly classified. We merged the 11 different walk directions of CASIA-B Dataset into 5 discrete directions as done in [10]. The confusion matrix computed is given by Table 3. Note that the average of classification rate (Balanced accuracy) 91.66% surpasses the rate obtained by [10] (90.87%) for a delay equal to 20 frames. With a delay of 40 frames, we reach the rate 97.15%.

We give in Table 4 the confusion matrix computed for Casia-A. The average of classification rate (Balanced accuracy) 96.14% surpasses the rate obtained by [10] (94.58%) for a delay equal to 20 frames. With a delay of 30 frames, we reach the rate 98.11%.

Results for Campus dataset

Once silhouettes are extracted (see Figure 9), their top and bottom points are located and the direction is derived. The error of computed direction depends on the segmentation errors and on the number of used silhouettes. Note that the more numerous are the silhouettes, the lower is the error (see Figure 9).

Our method takes into account the direction change as described in the proposed algorithm. New fitted lines are considered once the new top and bottom points are far from the previous fitted lines (see Figure 10).

Table 5 indicates the mean error and standard deviation depending on the delay (past frames) required to obtain an accurate direction estimation. Note that with a delay of 30 frames, the direction is estimated with accuracy.

Results for MGP01 dataset

The same experiments have been conducted on MGP01 dataset. Figure 11 shows that despite the presence of the shadow, our method delivers the correct direction for pedestrian whatever their orientation. Mean error and standard deviation are given by Table 6.

Table 4. Confusion matrix for Casia-B dataset with delay 0, 20, 30 frames. Values given are percentages.

Directions (delay)	D1	D2	D3	Accuracy	Accuracy [7]
D1(0 f)	79.73	15.94	4.32	79.73	92.5
D1(20 f)	89.18	10.24	0.58	89.18	
D1(30 f)	94.44	5.53	0.02	94.44	
D2(0 f)	0.52	93.22	6.25	93.22	92.5
D2(20 f)	0	99,56	0.04	99.56	
D2(30 f)	0	99,91	0.09	99.91	
D3(0 f)	1.22	7.26	91.51	91.51	98.75
D3(20 f)	0	0.3	99.70	99.70	
D3(30 f)	0	0	100	100	
Our(0 f)	Balanced Accuracy			88.15	94.58
Our(20 f)				96.14	
Our(30 f)				98.11	

Efficiency of the method

Time processing

The average time processing per frame estimated on used datasets is 16 ms. This means that for one second, we can process 62 frames which is in accordance with real-time applications. This time correspond to a limited number of operations required for lines fitting using the top and bottom points and the computation of the angle.

Note that this provided time processing is estimated without any optimization of source code.

Robustness to segmentation

If we note segmentation error as a small displacement of the top or bottom point of the silhouette from its correct position in images, then our algorithm takes into account this error because this alteration do not influence the lines fitting process. Despite if there is segmentation error, such as occlusion, part missing as illustrated by Figure 12, the error of estimated direction remains stable.

This robustness is constrained by the presence of the top or bottom point displacement in discontinued frames such as shown in the same figure.

We note also that our method estimates with accuracy the walking direction of pedestrians as we seen in the conducted experiment on Campus and MGP01 datasets. The processed segmented frames present many errors of segmentation such as missing or additional pixels of pedestrian's silhouettes due to noise and remained shadow despite the application of the method [15] for shadow elimination.

Failure cases

As we seen in the experiment section, our method needs a delay of at least 20 frames in order to deliver a correct estimate of the direction. The number of 20 frames represents the required sets of successive top and bottom points in order to obtain accurate fitted lines.

Our method fails when the number of successive frames including segmentation errors are successive and numerous. In this case, the slope of fitted line changes drastically and then the computed vanishing point provides inaccurate estimated direction. Figure 13 illustrates some successive frames from MGP01 dataset presenting missing parts. Once the top point of each new frame is far from the previous fitted line, a new line is fitted starting from that point, consequently, the computed vanishing point is inaccurate.

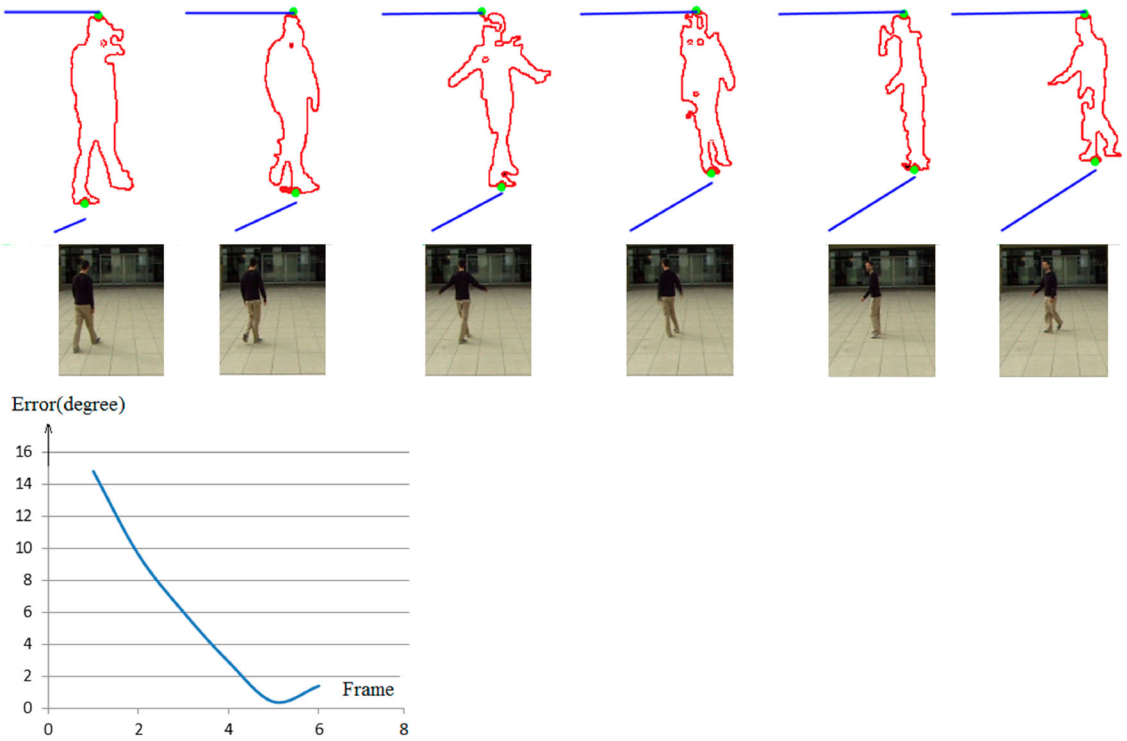

Figure 9. (Top) Frames from Campus dataset of the **same direction**. (Bottom) Variation graph of the error of estimated direction.

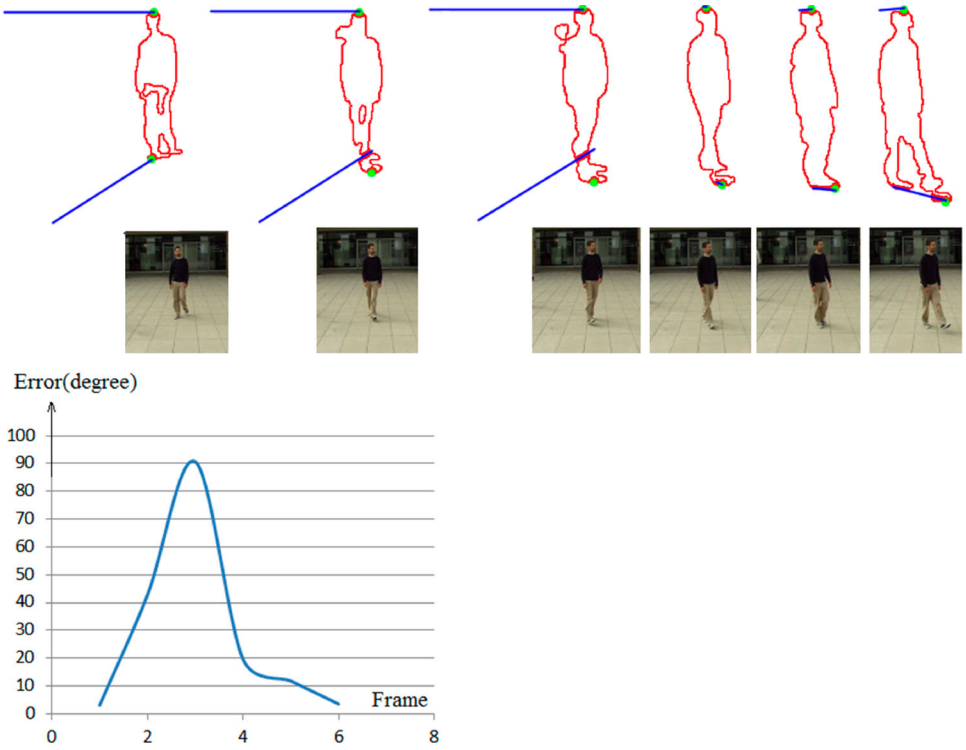

Figure 10. (Top) Frames from Campus dataset during **direction change**. (Bottom) Error variation of estimated direction.

Table 5. The mean error μ and standard deviation σ for Campus dataset related to the delay.

Delay	0	10	20	30	40
μ	14.9	10.9	7.7	4.7	3.5
σ	18.2	13.2	9.4	4.4	1.9

Table 6. The mean error μ and standard deviation σ for MGP01 dataset related to the delay (case of single pedestrian).

Delay	0	10	20	30	40
μ	12.51	11.52	9.1	7.67	6.7
σ	11.1	10.86	7.78	3.9	3.02

Case of failure occurs also when single or groups of pedestrians are grouped and constitute the same blob. In case where they are walking in different directions, the top or bottom points will be not visible for each group and the located ones may not define in 3D parallel lines as indicated by Figure 14.

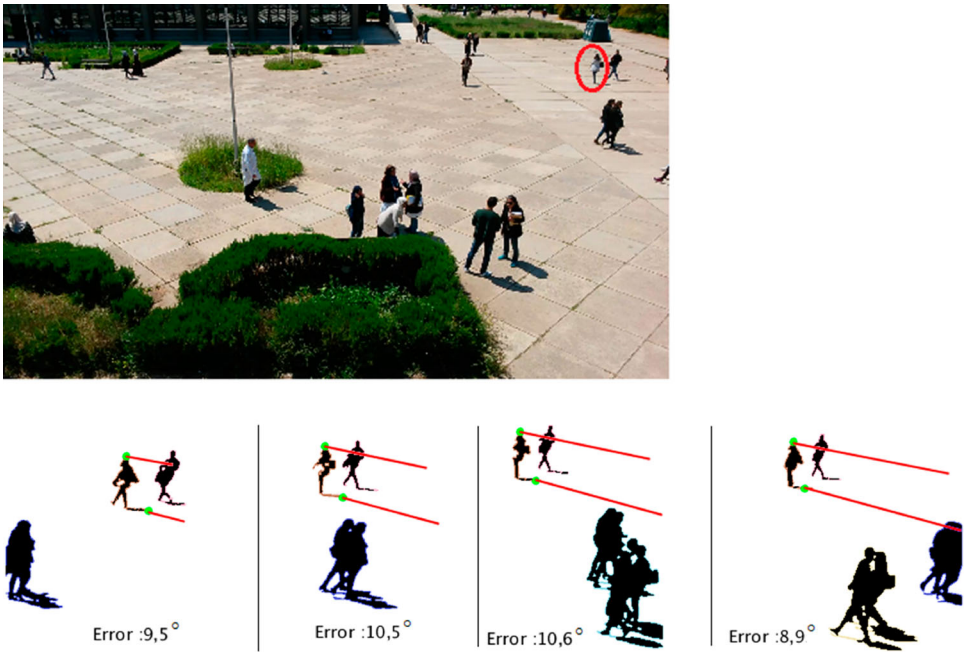

Figure 11. (Top) Frame from MGP01 dataset. (Bottom) Fitted lines and errors of estimated direction for encircled pedestrian in previous and next frames of the shown one in top. Source: [17].

Figure 12. Direction computation of walking pedestrian for successive frames of Casia-B dataset (Ground truth direction: 18°). Note that despite the missing of head in some frames, the error of estimated direction remains stable (5°).

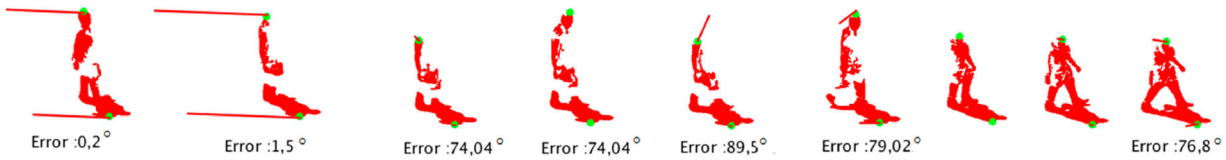

Figure 13. Example of failure which occurs processing successive frames contain segmentation errors. We note that the missing of head in the third frame, new fitting lines are generated, but the head appears in the next one, and so on.

Figure 14. Example of failure when groups of pedestrians are walking in different directions are grouped (case of second, third and fourth frame).

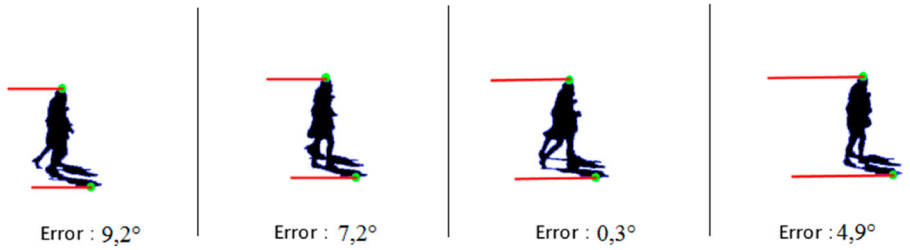

Figure 15. Example group of pedestrians walking in the same direction but maintaining the shape's blob invariance.

However, if the two pedestrians are in the same blob and moving in the same direction without change of the blob's shape, then the top and bottom points define two parallel lines in 3D and the vanishing point gives a correct direction as shown by Figure 15.

If the shadow is not removed from the pedestrian's blob or pedestrian group's blob and there is a variation of its shape. In this case, the set of bottom points located in images do not correspond aligned points in 3D. Consequently, the error of estimated direction

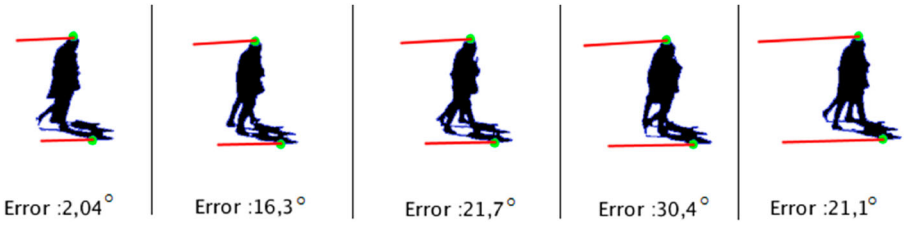

Figure 16. Example group of pedestrians walking in the same direction with some dispersion of shape's blob.

increases and is related to the amount of shadow dispersion. Figure 16 illustrates an example where the error increases because the shadow is not removed and its shape varies along frames.

Conclusion

Starting from a video sequence of walking pedestrians, we propose a new method for direction estimation with accuracy. We need a certain delay to provide this direction because our method is based on the fitting lines of top and bottom of moving blob in image.

Compared to the state-of-the-art, we surpass the proposed approaches working on single pedestrians. Note that estimating direction pedestrian group has not received yet attention from researchers; we plan to apply our approach for pedestrian group direction estimation in worst conditions for image segmentation.

In other side, our approach can serve for real-time applications because they need to perform a small number of operations.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Amina Bensebaa received the master degree in Intelligent Informatics Systems from the University of Sciences and Technology Houari Boumediene, Algeria, in 2012. She is also a PhD student at the University of Sciences and Technology Houari Boumediene, Algeria, under Professor Larabi's supervision. Her current research interests concern image and video analysis, content-based image retrieval and pattern recognition.

Slimane Larabi received his PhD in Computer Science from the National Institute Polytechnic of Toulouse, France, 1991. In January 1992, he joined the Computer Science Department of USTHB University in Algeria, where he is currently a Professor. He leads research in Computer Vision Group of the Laboratory of Artificial Intelligence Research. His work spans a range of topics in vision including: image description, human action recognition, head and body pose estimation and video analysis. He also proposed and leads the Master of Visual Computing in the same university and teaches several courses: data visualization, game design, multimedia systems, artificial intelligence and computer vision. He conducted many projects in different areas, especially in computer vision, augmented reality, and data visualization.

ORCID

Slimane Larabi http://orcid.org/0000-0001-8994-5980

Image notes

The left image used in figure 2 can be accessed through the following link: https://www.dezeen.com/2017/10/12/umbrellium-develops-interactive-road-crossing-that-only-appears-when-needed-technology/ and the right image has been taken from ref [2].

The image dataset used for figures 3 & 4 can be accessed through the following links
https://towardsdatascience.com/
http://www.cvg.reading.ac.uk/PETS2006/data.html

The images in figure 8 (left) and figure 11 were taken from the dataset MGP01 which can be accessed through the following link: http://perso.usthb.dz/~slarabi/MGP01.html and the right image of figure 8 can be accessed through the following link: https://cvlab.epfl.ch/data/data-pom-index-php/

References

[1] Mairs J. Umbrellium develops interactive road crossing that only appears when needed, Online, 2017, 2018-02-3. [Online]. https://www.dezeen.com.
[2] Kooij JFP, Schneider N, Flohr F, et al. An example of pedestrian intention estimation using contextual cues. Context-based pedestrian path prediction, in European Conference on Computer Vision (ECCV), 2014, pp. 618–633.
[3] Teaching cars to see vehicle detection using machine learning and computer vision. https://towardsdatascience.com/.
[4] Shimizu H, Poggio T. Direction estimation of pedestrian from multiple still images., IV 2004.
[5] Guangzhen Z, Mrutani T, Kajita S, et al. Video based estimation of pedestrian walking direction for pedestrian protection system. J Electron (China). 2012;29(1/2).
[6] Baltieri D, Vezzani R, Cucchira R. People orientation recognition by mixtures of wrapped distributions on random trees, In proc. Eur. Conf. Comput. Vis. (ECCV), 2012, pp. 270–283.
[7] Liu H, Ma L. Online person orientation estimation based on classifier update, In Proceeding of IEEE Int. Conf. Image Processing ICIP, 2015, pp. 1568–1572.
[8] Tao J, Klette R. Integrated pedestrian and direction classification using a random decision forest, ICCV Workshop, 2013.

[9] Yu S, Tan D, Tan T. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition, ICPR, Hong Kong, China. August, 2006.
[10] Raman R, Sa PK, Majhi B, et al. Direction estimation for pedestrian monitoring system in smart cities: An HMM based approach. IEEE Access. 2016;4:5788–5808.
[11] Raza M, Chen Z, Rehman S, et al. Appearance based pedestrians head pose and body orientation estimation using deep learning. Neurocomputing. 2018;272:647–659.
[12] Huang S, Ying X, Rong J, et al. Camera Calibration from Periodic Motion of a Pedestrian, CVPR 2016.
[13] Junejo IN. Using pedestrians walking on uneven terrains for camera calibration. Mach Vis Appl. 2011;22(1):137–144.
[14] Guan J, Deboeverie F, Slembrouck M, et al. Extrinsic calibration of camera networks based on pedestrians. Sensors. 2016;16:654. doi:10.3390/s16050654.
[15] Cucchiara R, Grana C, Piccardi M, et al. Improving Shadow Suppression in Moving Object Detection with HSV Color Information, IEEE Intelligent Transportation Systems Conference Proceedings - Oakland (CA), USA - August 25–29, 2001.
[16] PETS2006. http://www.cvg.reading.ac.uk/PETS2006/data.html.
[17] Lv F, Zhao T, Nevatia R. Self-calibration of a camera from video of a walking human, IEEE International Conference of Pattern Recognition, 2002, pp. 281–304.
[18] MGP01. http://perso.usthb.dz/~slarabi/MGP01.html.
[19] Berclaz J, Fleuret F, Turetken E, et al. Multiple object tracking using K-shortest paths optimization. IEEE Trans Pattern Anal Mach Intell. 2011;33(9):1806–1819.
[20] Fleuret F, Berclaz J, Lengagne R, et al. Multi-camera people tracking with a probabilistic occupancy map. IEEE Trans Pattern Anal Mach Intell. 2008;30(2):267–282.