Artificial Vision for Road Safety Improvement

Enrique Cabello Pardos (1*), M. Araceli Sánchez Sánchez (2), Laura Agudo Mérida (1), Antonio Sanz Montemayor (1), Iñigo Martín Sánchez (2).

(1) Universidad Rey Juan Carlos
ESCET
C/ Tulipán s/n
28933 Móstoles (Madrid)
Spain

(2) Departamento de Informática y Automática
Universidad de Salamanca
Plaza de la Merced s/n
37008 Salamanca
Spain

***Corresponding author:**
Tel.: +34 – 91 – 664 74 55 FAX: +34 – 91 – 664 74 90 Email: ecabello@escet.urjc.es

ABSTRACT

This paper presents a developing project applying image processing algorithms and human behaviour techniques to study pedestrian-vehicle conflicts. Real data will be obtained from two zebra crossings in the city of Salamanca (Spain). Results will be a set of measurable parameters influencing the conflicts. Image processing techniques allow us to manage huge and real information data. One overview and two lateral cameras will give us videos to be processed.

I.- INTRODUCTION

A dramatic problem in modern cities is the huge number of injured and dead persons by traffic accidents. Road accidents have been ranked as the eight largest cause of death in the world (World Health Organisation 1996). Over 45.000 people are killed on the roads in the EU every year, and there are 1.5 million reported casualties. This figure could be as high as 3.5 million when under-reporting is taking into account. More than a half of these accidents occur in urban areas, with a disproportionately large number of vulnerable road users being killed by cars. Nearly half the deaths in urban areas are now outside the vehicle. Even though the total amount of fatalities has decreased since the early 1990s in many European countries, the ratio of fatal to all injury accidents in urban areas still remains a major concern.

In this scenario, several research projects were promoted, mainly in the USA, Canada and Australia and also in Europe, to study traffic accidents. We can consider DUMAS project, in the EU 4 Framework Programme [1], as a pioneer study in Europe. One of the working packages developed by the consortium was the study of accidents in different European cities. In this project were reported that the lack of a standard definition of accident is a major issue to be solved, also the circumstances influencing an accident were not studied.

The work presented in this paper promotes a systematic study of conflicts between pedestrians and vehicles and as a final conclusion, a set of measurable parameters influencing a conflict will be obtained. The main goal of our research is the study of conflicts. In zebra crossing more conflicts that accidents are present. All accidents are preceded by a conflict, even not all conflicts generates an accident. Also, statistically significant data could be obtained from conflicts and no social repulse is involved studying conflicts.

Till now, only police, assurance or hospital reports have been considered as a evidence source in the parameters involved in an accident. These reports have numerous drawbacks due to the absence of neutral evidences. Even if witness exists, contradictions are quite common. To cover this gap, the Spanish Traffic Directorate (DGT) using technology now available has promoted a systematic study. This technology is based in computer vision and in high level digital processing. But this is only one side of the problem, other side is the study of the circumstances affecting human behaviour. Our work has been split also in two parts, one covers computer science development and the other will cover psychologist factors. From a computer science point of view, we could record the images taken in real life to be studied later. This study can be done in real time or off-line. In the first case, part of the information will be lost, second gives us the opportunity to isolate all the factors and to compute it with enough accuracy to see how affects in conflicts.

Our work will study a zebra crossing with traffic light during 15 days, during other 15 days a zebra crossing without traffic light will be studied. The selected city for the study is Salamanca (Spain). With the help of a computer science team and a psychologist team both sides of the study will be covered.

Present paper will show initial results, conclusions and discussions because project is now on-going and has not finished. But we think that could be useful to show lessons learned and how this kind of projects could promote the use of available technology to obtain better cities, or at least more safety.

The structure of the rest of the paper is as follows. Section II provides a short overview of previous research. Section III is devoted to show a project description following by Section IV devoted to the technical set-up of our system and in Section V we included a detailed explanation of the algorithms implemented. Section VI presents actual and expected results. We end with conclusions, lessons learned and future work in Section VII.

II.- STATE OF THE ART

Pioneer European project DUMAS [1] proposes solutions for the Urban Safety Management, studying high-risk crossings in cities and accidents produced there. DUMAS consortium takes into account public transport and measurements of vehicles and pedestrians in a statistical way and they propose law enforcement to improve security.

In parallel way, artificial vision has proven to be a useful tool that can be applied in topics related with urban safety and traffic. Literature on motion detection and multiple object tracking over video image processing is rather extensive and has been applied in traffic situations [2,4,7,8,9,11,13, 16,17,18,19]. Simplest application is the detection of critical situations, as congestion [19], and focusing the human operator’s attention on the problem. In such cases, human supervision is still considered unavoidable.

New algorithm is proposed in [16] to improve the detection and tracking of pedestrians. It reduces the number of active regions detected and tries to avoid occlusions and unreliable associations. Main drawback is that longer time is required for analysis and can not track pedestrians due to merging with other objects. Likewise, [17] described an application that deal with the monitoring and measurement of vehicles and pedestrians traffic within road or public transport systems.

Most problems in artificial vision can be solved assuming some kind of a priori knowledge of the concrete scene. This “closed-world” is often used to reduce the redundant information and so, increasing the speed of the algorithms, but these methods are highly tied to concrete environment. Such “closed-world” assumptions are proposed for tracking vehicles in traffic scenes [13,16,17] where the ground-plane constraint reduces the problem of localisation and recognition as much as we all know that road vehicles stand on the ground-plane. But in [18] high-level computer vision algorithm based in fuzzy logic is presented. The goal of the algorithm is to classify each pixel of the frame in one of four classes (Moving, Still, Covering Background and Uncovering background). This classification permits to shape more precisely the objects boundaries and to estimate vehicles speed.

Some papers presents a pre-processing step to reduce noise (by median filters [2], time convolution filters,...), others do not [4,6]. Once this is done the segmentation part is taken into account. Edge and local feature detection [5], extraction of boundaries, lines and curves [9], morphological processes such as dilation and erosion [2], background subtraction [4, 12, 14], other motion detection techniques (gradient based [2,6], optical flow [5],...) and a priori scene knowledge: ground-plane [13,16,17], white lines in a road [9,11],... are most valuables methods. Template matching is often used when the objects have a well-known shape. T.N. Tan et al [13] reduce the number of degrees of freedom using ground-plane constraint and then, applying template matching to identify best pose.

As shown in [3], tracking techniques as correlation tracking where a template is extracted from the first image and is compared with the objects in next frames could be used. Main problem is that it is very sensitive to occlusion, light conditions and changing size objects. Motion tracking, mainly optical flow methods, are also quite sensitive to occlusion due to its dependence on brightness smooth changes. In order to avoid the size changing restriction of the correlation tracking, deformable template matching permits small, constrained, changes in the template over time. It is usually used when the template matches an object, which is moving along a perspective trajectory, and so it has a geometric distortion between consecutive frames. Kalman filters is another common tracking method used by authors [4,8,10,11]. As [15] indicates “this filter is a set of mathematical equations that provides an efficient computational (recursive) solution of the least squares method”.

III.- PROJECT OVERVIEW

Aim of this project is study of zebra crossings and risky situations involving pedestrians and vehicles whether traffic lights are presented or not. Real pedestrian crossings will be studied in the city of Salamanca (Spain). This is a key point because real data will be important to obtain inferences based in empirical data. In this project two collaborative teams are presents, one computer science group, responsible of image processing algorithms, and a psychologist group in order to study human behaviour in conflicts between pedestrian and vehicle. Basic scheme of the project is shown in Figure 1.

Image processing task is to record sequences in zebra crossings. Only conflict sequences will be studied to do this, an automatic detection and prediction system of conflicts has to be built. The initial phase will end with an objective and automatic detector and recognition system. Later, psychologist team has to define the physical variables (parameters) to formulate a conflict definition. With this definition and the variables extracted from the zebra crossings conflicts will be analysed in an unsupervised way giving answers to the questions proposed by the DGT about how to avoid this kind of accidents.

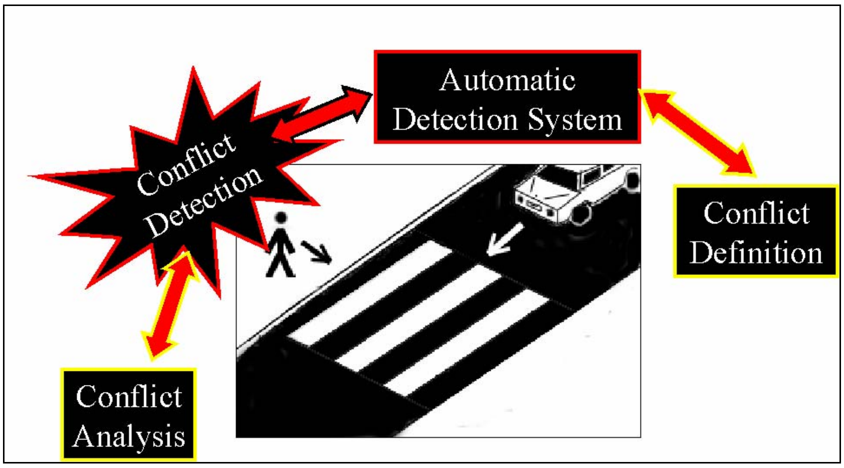

Figure 1: Basic scene with project overview description

IV.- TECHNICAL SET-UP

Although this project will cover a long period of time (more than 240 recording hours) comparing with similar works, this is only a small fraction of the time daily spent by humans in such scenes. To provide information really used, two dangerous zebra crossing in the city of Salamanca (Spain) were selected to be studied (see Figure 2). This selection was conditioned by other constraints, due mainly to the set-up needed; one extremely important factor is visibility, to avoid occlusions due to trees, buildings or any other city elements.

Image acquisition is done by three small digital cameras (see Figure 3) although most important is an overview one. The other two cameras are placed laterally to take scene details, such as cars parked close to the crossing, the pedestrian attention and behaviour. Once the overview camera detects a movement, information extracted from lateral cameras will be important for the psychologist group. However, parameters will be measured from overview camera.

Figure 2.- Selected zebra crossings in Salamanca (Spain).

Work has been done using wireless cameras transmitting compressed images using radio frequencies (2.4 GHz). These cameras need a computer to be located at 100 meters maximum with radio-receptors. This solution offers us enough flexibility to place cameras in small boxes and placing a computer in a secure place. Computer is a high performing computer to record images.

One key point in human behaviour experiments is that the observer could not be observed. This constrain is extremely important in our situation. Pedestrians and vehicles has to maintain its normal conduct, otherwise the experiment will be unsuccessful. This could be achieved easily with small camera containers.

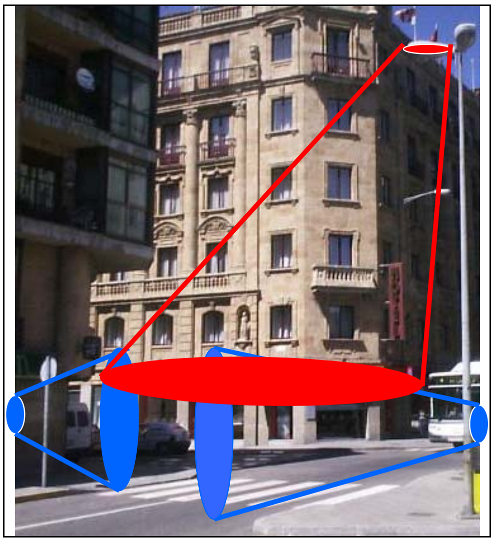

Figure3: Camera placement scheme.

V.- ALGORITHM DESCRIPTION

Software presented in this section corresponds to the automatic detection and tracking of vehicles and pedestrians. As explained later, usefully computer vision information to measure parameters will be obtained from overview camera, so only this camera will be considered later. Human referees have done information extraction from lateral cameras. The only consideration is the three cameras have to considered the same time stamp. Overview camera algorithm has been structured in two phases depending on the treatment applied to the images: low level image processing (to obtain a set of candidate pedestrian and vehicles in the image) and high level image processing (using an artificial neural network and a tracking algorithm to ensure correct detection). Figure 4 shows an overview of the decomposition of the algorithm. Each phase is later described in detail.

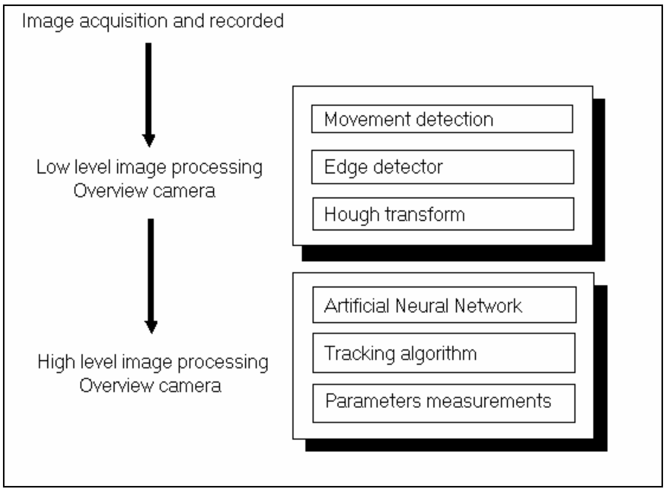

Figure 4. Scheme of the algorithm.

A. LOW LEVEL IMAGE PROCESSING

Preprocessing (low level image processing) is a critical step in which input is a colour image and final output is a set of windows in the image that may contain vehicles or pedestrians. Preprocessing algorithm has to deal with all the problems related to noise and colour and hence a robust algorithm has to be built. This algorithm is split in several steps. Each one represents the application of a process in order to isolate an object in the image. Some of the algorithms considered are standard but others have been tailored to our problem.

Step 1: Motion detection is performed to obtain moving objects in the image. To perform this task a classical optical flow is applied. Motion detection algorithm operates with two frames applying optical flow to the three-colour components RGB (Red - Green - Blue). We set a threshold to avoid noise detection, this threshold depends on the input frame. Those pixels with optical flow value greater than the threshold are represent in blue colour (fig.2) and are called candidate pixels, everyone else is set to white. Once the full image is analysed, only regions with enough blue pixels are considered. These regions will be the ones processed late and the rest of the image can be rejected without loosing significant information, because small blue aggregations of pixels could be due to noise or small objects. Those extracted windows or regions of interest (ROI) contain the supposed moving object. If occlusion is present and two or more ROIs overlapped, we integrate all the ROIs into a single one.

Step 2: Several edge detectors (Roberts, Prewitt and Sobel) have been applied best one for our purposes was 3x3 Sobel edge detector. Edge detector output is the input required by Hough Transform. Hough transform will be used to obtain straight lines in candidate regions. Vehicles are formed by edges, so an accumulation of edges gives us a clue that a vehicle is present.

For recognition of well-defined linear objects, such as cars or vehicles in general, Hough Transform (using polar coordinates to avoid the difficulties that it presents in vertical line detection) is proposed. It is a heavy process so more efficient alternative method (neural nets, PCA,...) will be considered.

B. HIGH LEVEL IMAGE PROCESSING

Artificial neural networks are a well-known classifying technique and have repeatedly shown its potential as a tool for classification purposes [20][21][22]. A neural network has to decide if a moving object can actually be classified as vehicle, pedestrian or not: the input is a gray level image and output is its classification as vehicle, pedestrian or unknown

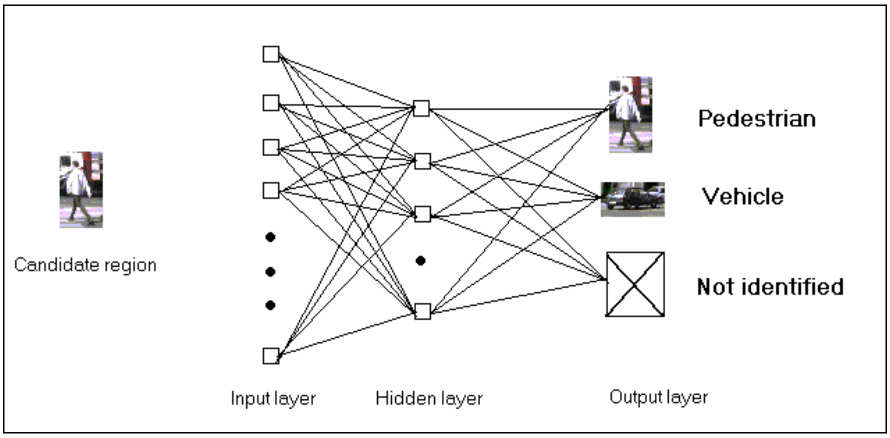

Figure 5.- Artificial Neural Network scheme.

Recursive tracker is based on the Kalman filter model that is known to be an optimal linear estimator. For each object, at each time, the filter maintains an estimated state as well as its spatial extent and confidence factor. Given a target at a determinate time, prediction equations predict its new position in the next frames. Prediction equations are linear estimation of movement based on a Taylor series approximation. The "order" of the process is determined by the ratio of processing rate compared to the rate of expected change of the (N+1)-th derivative. In our case targets undergo very high acceleration or deceleration, so acceleration, velocity and displacement has to be estimated.

Once the image has been processed, physical parameters (speed, distance between vehicles and pedestrians, pedestrians groups, and so) could be extracted from frames. This is the final step for automatic conflict detection. Then, one a conflict is detected, frames from the three images will be shown to human referees.

VI.- ACTUAL AND EXPECTED RESULTS.

As an ongoing project, only preliminary results could be shown but good results in movement detection has been achieved. Optical flow algorithm works well with testing videos, which are worse than definitive ones. Relatively small cars in a highway are detected from a camera that is not overview placed so perspective distortion is considered and sometimes occlusion appears. Those regions of interest that overlapped are linked to form one, preventing cuts in object.

Figure 6. Initial frames of our experimental videos

Figure 7. Detection of moving objects.

We are focused now in next step in image processing, recognition of moving objects in scene. For this stage, regions took from the movement detection procedure must have the same sizes as input for artificial neural network. Tracking method is now implemented so next task is movement prediction.

VII.- CONCLUSIONS. LESSONS LEARNED AND FUTURE WORK.

One of most important thing learned, is that co-operative work between different institutions, local government of Salamanca, Universidad de Salamanca, Universidad Rey Juan Carlos and Spanish Traffic Directorate (DGT: Dirección General de Tráfico), is needed to face a very complex problem, like the study of vehicle and pedestrian conflicts. Also, only one expert group (psychologist or computer vision) could not cover the entire work. Partnership is also needed between experts to obtain not only reliable solutions but also the parameters affecting human behaviour.

In our work it had been unavoidable to keep in privacy the identity of the people who appears in the videotape. Only if people do not know that are being recorded, they will act naturally. This was one major problem in order to obtain official permissions. Aim of the project is not to identify or punish bad conduct. The objective is to obtain information about behaviours that causes conflicts and accidents.

Future work will be directed in two ways. First way is to obtain better algorithms, faster and more accuracy. In this way we would like to obtain real time image processing without loosing significant data. Second direction is to produce modifications on pedestrians crossing that can decrease the number of casualties. Finally, it will be possible to inform pedestrians and drivers about which behaviours cause accidents.

ACKNOWLEDGMENTS

This paper presents preliminary results in a project granted by the Spanish Traffic Directorate (DGT: Dirección General de Tráfico) and the Universidad de Salamanca, both supports are kindly appreciate.

REFERENCES

[1] DUMAS Proyect, Final Report. EU 4^{TH} Framework Programme. Contract Nº RO-96-SC.201.

[2] D.J. Dailey, L. Li, "An Algorithm to Estimate Vehicle Speed Using Un-Calibrated Cameras"

[3] S. S. Intille, "Tracking Using a Local Closed-World Asumption: Tracking in The Football Domain", M.I.T. Media Lab Perceptual Computing Group Techn. Report No. 296

[4] D. Koller, J. Weber and J. Malik, "Robust Multiple Car Tracking with Occlusion Reasoning", In Proc. European Conf. Comp. Vis., Vol 1, Pp 189-196, Stockholm, Sweden, May 1994

[5] Berthold K.P. Horn and Brian G. Schunck, "Determining Optical Flow", Artificial Intelligence, volume 17, pages 185-204, May 1981.

[6] J. Amat, A. Casals and M. Frigola, "A Specialized Vision System for Control by means of Gestures", Advanced Control Workshop, Coimbra, pages 678-683, 1998

[7] J. Badenas, F. Pla, "Segmentation Based on Region-Tracking in Image Sequences for Traffic Monitoring", In 14th International Conference on Pattern Recognition, ICPR, 1998.
[8] R. García, J.Batlle, Ll. Magí, Ll. Pacheco, "Seguimiento de Múltiples Objetos: Un Enfoque Predictivo"
[9] J.M. Sanchiz, R. Pérez, M. Aguilella, "Interpretación de Escenas de Tráfico para Asistencia a la Conducción"
[10] S. Watcher and H. H. Nagel, "Tracking Persons in Monocular Image Sequences", Computer Vision And Image Understanding, volume 74, pages 174-192, June 1999.
[11] W. Kasprzak and H. Niemann, "Adaptative Road Recognition and Ego-state Tracking in the Presence of Obstacles", Int. Journal Computer Vision, vol. 28, pages 5-26, 1998.
[12] Stephen S. Intille, James W. Davis, and Aaron F. Bobick, "Real-Time Closed-World Tracking" IEEE Conf. on Computer Vision and Pattern Recognition, November 1996.
[13] T. N. Tan, G. D. Sullivan, and K. D. Baker, "Model-Based Localisation and Recognition of Road Vehicles", Int. Journal Computer Vision, vol 27, pages 5-25, 1998.
[14] Y. Ivanov, A. Bobick and J. Liu, "Fast Lighting Independent Background Subtraction" M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 437.
[15] G. Welch and G. Bishop, "An Introduction to the Kalman Filter", UNC-Chappel Hill, TR 95-041, Nov 2000.
[16] A. Anzalone A. Machi, "Video-based management of traffic light at pedestrian road crossing", Advanced video-based surveillance systems, Pps 49-57, Kluwer Publ. 1999.
[17] Jean-Marc Blosseville, "Image procesing for traffic Management", Advanced video-based surveillance systems, pages 67-75, Kluwer Academic Publishers 1999.
[18] M.Barni, F. Bartolini, V. Cappellini, F. Lambardi and A. Piva, "Fuzzy motion detection for highway traffic control", Advanced video-based surveillance systems, pages 58-66, Kluwer Academic Publishers 1999.
[19] M. Pellegrini, P. Tonani, "Highway traffic monitoring: main problems and current solutions", Advanced video-based surveillance systems, Pps 27-33, Kluwer Publs 1999.
[20] B. D. Ripley. Pattern recognition and neural networks. Cambridge Univ. Press. 1996.
[21] C. M. Bishop. Neural networks for pattern recognition. Oxford University Press.1995
[22] C. G. Looney. Pattern recognition using neural networks. Oxford Univ. Press. 1997.