Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 46(2), 2019, Pages 366–371
ISSN: 1223-6934

Object detection by grouping partial perception phanges

Nicolae Constantinescu

Abstract. This paper deals with exhibiting and classifying objects along an image sequence. Motion is considered through the image sequence, leading to a frame by frame analysis. We first review a set of methods to detect mobile objects, knowing that our intended perception model is to include colour and get robust the acquisition of flickering images. Objects are being detected due to the visual variances of observer's perceptions: different segments of the image string are compared. The core of the image-analysis process deals with defining methods to reveal object structures. These methods piece together tasks of detecting new objects, grouping them, identifying the objects distribution along the scene. All contribute to eliminating collateral information.

2010 Mathematics Subject Classification. Primary 65S05; Secondary 65S99.
Key words and phrases. image analysis, visual perception, histogram modes, object detection.

1. Introduction

The end goal of the study described in this paper is to discard uninteresting objects efficiently from a dynamic scene. An image is classically defined here as an array of points $C_{(i,j)}$, where a colour measure C is assigned to every position (i,j). A region in this image is a 4-connected set of pixels showing a given common property. In Image Processing many region extraction techniques use some statistical analysis, ranging from mere histogram identification to sophisticated clustering, to set relations between frame subparts. These techniques may thus be associated to progressive merge from pixel up to connected components or to tentative straight component extraction based on splitting techniques from image down to components. A dynamic frame is numerically a sequence of time-indexed static frames, giving access to information about the object's dynamics. Here again the trace of image parts can stem from a statistical analysis of the image $-$ or image sub-sequence $-$ data e.g. based on the associated histograms. After the article by J.L. Muerle and D.C. Allen ([1]), many approaches were proposed based on block segmentation and histogram analysis to raise synthetic information on grey levels (according to [2, 4]). Unlike such classic approaches, in this article we foster a combined colour nuance detection mechanism which applies better in the case of flicker. The aim of this research part is to define object classes and for each newly found structure to decide whether it is part of an already defined class or it initiates a new class to be defined. For an existing class, the detected structure should be provided with a probability that it already belongs to the object database or it is a new object in that class.

Received March 19, 2019. Accepted September 26, 2019.

2. Describing the study

Our approach consists of splitting each and every image of the sequence into regions. Each region has attached a certain set of informations. Based on this information, objects in the frame are exhibited and analysed following a scheme similar to the one by Sharat Chandran and Naga Kiran ([5]). Transformed histograms are first generated for each part of the image. They are then compared aiming to identify the transformation between images translating the dynamics in their sequence. Comparing is achieved by measuring the correlation degree between the transformed histograms computed on blocks at similar positions along the sequence $(1..n)$ and with the same in image $n+1$ (for details, see [6]). Based on the correlation degrees, the system decides the zones where the new structures belong. These structures become candidates to be matched with a new object. Through this process, new informational structures associated with objects are dynamically added. A few properties of theirs are instantiated along the process, further contributing to object comparison and classification. The latter method thus achieves the conversion of observed changing zones into a classification of informational structures that appeared along a passed picture sequence, with reference to the current picture (or a set of most recent pictures). Afterwards, informational classes will be defined, so called objects, thanks to an object comparison function. Let be $(I_1, I_2, ..., I_n)$, a picture sequence acquired from an observation point. We consider colour images unlike classical approaches, Mohan M. Trivedi and Charles A. Harlow and Richard W. Conners and Semoon article ([8]). $I_{(n+1)}$ is the last image, from the same view-point, Sameer Singh and Markos Markou article ([9]).

3. Analysis and proposed method

3.1. Model construction. Let be I_1 the first image of the first frame. To this image will be applied both a geometric scale change transformation and a colour translation of the initial palette into a colour scale adapted to the processing power of the system (according to Nikoletta Bassiou and Constantine Kotropoulos article, [10]). $h_{(rez)} \times v_{(rez)}$ will be the new resolution and h_{deep} the colour depth which the image is coded with. The image is divided into analysis cells, obtained by horizontal and vertical image splitting, like in Figure 1.

In this way, the image will be made of $m \times n$ rectangular neighbour cells, marked as

$$\theta^1_{(i,j)}, (1 \leq i \leq m, 1 \leq j \leq n) \tag{1}$$

We define the histograms bound to zones (i,j) as $h^1_{(i,j)}, (1 \leq i \leq m, 1 \leq j \leq n)$, obtained by averaging frequencies, and also the associated histogram $LRS : H^1_{(i,j)}, (1 \leq i \leq m, 1 \leq j \leq n)$. Each histogram $h^1_{(i,j)}, (1 \leq i \leq m, 1 \leq j \leq n)$ brings out a transformed histogram, $h^t_{(i,j)}$, by combining histogram classes in the transformation function Θ, to decrease the informational noise: the collateral information is reduced in flattening the values which have a large dispersion. The object's database is empty at start.

FIGURE 1. Sppliting scheme.

3.2. Parameter computation. The basic histogram and the even-minded histogram are computed for colour l according to

$$h^t_{\theta_{(i,j)}}(l) = \sum_{k_1=1}^{hs} \sum_{k_2=1}^{vs} \omega_{\theta_{(k_1,k_2)}}(l), \tag{2}$$

$$ht^t_{\theta_{(i,j)}}(l) = \frac{h^t_{\theta_{(i,j)}}(l)}{\sqrt{(\sum_{l=0}^{h_{deep}} h^t_{(i,j)}(l))/h_{deep}}} \tag{3}$$

where $\omega^l_{(k_1,k_2)}$ is 1 when the value of the colour from (k_1, k_2) is equal to l and to 0 otherwise. From here, the histogram associated with the $\theta_{(i,j)}$ cell gets the form $h^t_{\theta_{(i,j)}} = \left\{ h^t_{\theta_{(i,j)}}(l) \right\}_{0 \leq l \leq h_{deep}}$. This represents the collection of histograms for each colour from the associated palette $[0 - h_{deep}]$.

3.3. Analysing frames. Let be $\delta^{(t,t+1)}_{(i,j)}(l)$ the value of the frame-to-frame difference I_t / I_{t+1}, for the colour value l. A cell will be considered candidate to he next group-analysis step (according to Kazunori Onoguchi article, [13]), if it shows a colour value greater than ξ_1 with difference bigger than a value ξ_2. The correlation level will be defined for such cells, as

$$L^{(t+1)}_{(i,j)} = max \left(L^{(t+1)}_{1(i,j)}, L^{(t+1)}_{2(i,j)} \right), \tag{4}$$

where

$$L_1 = \frac{\sum_{d=0}^{h_{deep}} \left| \left(h_{(i,j)}^{(t+1)}\right)_d - \left(h_{(i,j)}^{(t)}\right)_d \right|}{\sqrt{\sum_{d=0}^{h_{deep}} \left(h_{(i,j)}^{(t+1)}\right)_d^2} \sqrt{\sum_{d=0}^{h_{deep}} \left(h_{(i,j)}^{(t)}\right)_d^2}}, \tag{5}$$

$$L_2 = \frac{\sum_{d=0}^{h_{deep}} \left| \left(h_{(i,j)}^{(t+1)}\right)_d - \left(h_{(i,j)}^{(t)}\right)_d \right|}{\sqrt{\sum_{d\delta} \left(h_{(i,j)}^{(t+1)}\right)_d^2} \sqrt{\sum_{d\delta} \left(h_{(i,j)}^{(t)}\right)_d^2}} \tag{6}$$

$d\delta$ is assumed to be a very small a number which represents the histogram change. A cell will belong to the group $\mathcal{M}$ (group of cells which makes objects by construction) if it is a candidate cell and the correlation level is higher than a value ξ_3. The values $\xi_{1,2,3}$ are preset from experience results.

3.4. Object management. Let be $\mathcal{M}$ an aggregation of cells as previously defined. For each cell, a grouping algorithm will be applied, to construct zones of associated cells. Clustering the cells is constrained by the adjacency rule. A zone of candidate neighbours is first defined for each cell. The optimal group where a cell belongs depends on the so called attraction-point rule. It leads to build clustered zones composed of both grouped cells and cells included in a minimal wrapping rectangle (Zisheng Cao and Feng Chen and Youtian Du, [14]). An histogram is assigned to each zone as follows: $h_{z_k}(l) = \sum_{v=1}^{p} h_{(i,j)}^{(t+1)}(l)_{(i,j)\in z_k}$ represents the histogram associated with the z_k zone for the l nuance in the color scale $[0 - h_{deep}]$. The histogram associated to a zone for the whole spectrum is thus defined as:$h_{z_k} = \{h_{z_k}(l)\} \quad l \in \{1, 2, \ldots h_{deep}\}$ The feature relevant to each zone is its associated histogram.

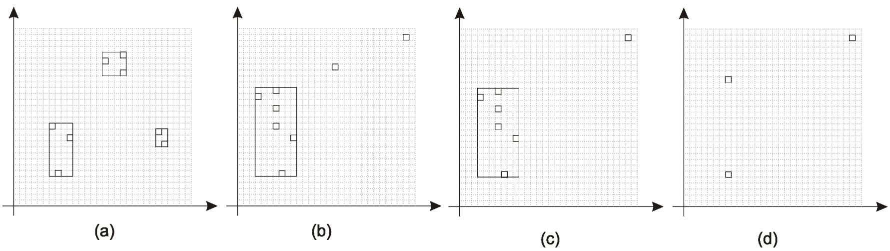

FIGURE 2. Grouping Cells. (a) When cells are well grouped the related regions are created without ambiguity. (b) Isolated cells: The decision to group such cells is taken after the previous frames, the current cell differences not being decisive enough. (c) Isolated Cell: In the case of only one isolated cell it is merely ignored. (d) Isolated cells only : in the case when the majority of cells are isolated the decision from the current frame comparison is taken after a number of previous frames, and delayed to next frame.

Throughout the process above objects from the database can be characterised accounting for the images of a video sequence. In this way, an $\mathcal{O}_i$ object is assigned to each zone. At each frame, an object may be added if the zone by which it is exhibited shows distances to already existing objects in the database greater than a value ξ_4 - again preset after experimental identification, which can applied to automate the studies from [7, 11]. The comparison between zones is performed in a similar way as of the cells, $\Theta(\zeta_{z_k}, L_{z_k})$, where L_{z_k} is the correlation function computed over the zone z_k.

4. Results

We implemented the techniques detailed above in a real time framework. A set of two dozen scenes was tested. Test sequences deal with flickered movies and the application extracts parts which are actually moving to construct an object structure out of them. The results can be shown in figure 3.

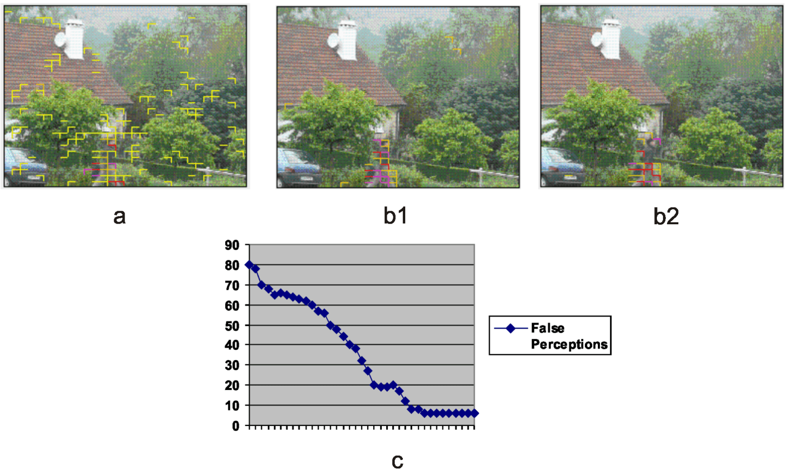

FIGURE 3. Object on the move. Flickered Perception Case. (a) Classical detection, based on the sole perception changes. (b1) Detection by our techniques. In the case of brief data (at this stage, the second part of the techniques do not get enough information to work, and the results show more "perceptions noise"). (b2) In the case of an object data base, the detection is finer and noise is reduced. (c) The evolution of the "false perception" number for the tested frames.

4.1. Case study. According to figure 3, when the noise detection is not active as in (a), any perception change is detected from a flickered source. By triggering the flicker detection on, we eliminate perceptions due to image flickering (b1). After a number of runs (denoted by $numbOfScans$) and object classes clustering, the frame by frame analysis improves the "true detection" to "false perception" ratio (b2). The number of runs, $numbOfScans$, depends by the quality of analysed image. In 3 (c) the evolution of the "false perception" number is sketched, this scheme can be applied to automate the studies from [3, 12]. It shows that, after a while - once the structure collection in the database is achieved to a satisfactory level for the given image and dynamics- it stabilises.

5. Conclusions

In this paper was presented a technique to detect and classify moving objects despite a flickered capture. Were combined histograms analysis for long and short frame-sequences, with perceptual zones classification. The novelty relies on the way to compare zones and the way colour is handled to eliminate "false perceptions" for the same structures. That way flicker splits from real moving objects for them to be detected. Based on the current results, the future work is to adapt our process from flicker to moving view angle. The end application is a real-time understanding of the object behaviour from the dynamic objects-classification in the case of a scene perceived by a moving car.

References

[1] J.L. Muerle, D.C. Allen, *Experimental evaluation of techniques for automatic segmentation of objects in a complex scene*, Pictorial Pattern Recognition, Thompson Book Co., 1968, 21–32.

[2] A.S. Abutableb, Automatic thresholding of gray-level pictures using two-dimensional entropy, *Comput. Vision Graph. Image Process.* **47** (1989), no. 1, 22–32.

[3] L. Duica, E. Antonescu, M. Pirlog, T. Purnichi, J. Szakacs, M. Totan, B. Vintila, M. Mitariu, C. Mitariu, A. Stetiu, Clinical and Biochemical Correlations of Aggression in Young Patients with Mental Disorders, *Revista de Chimie* **69** (2018), no. 6, 1544–1549.

[4] A. Nakib, H. Oulhadj, P. Siarry, Image histogram thresholding based on multiobjective optimization, *Signal Processing* **87** (2007), 2516–2534.

[5] S. Chandran, N. Kiran, Image retrieval with embedded region relationships, *SAC '03 Proceedings of the 2003 ACM symposium on Applied computing*, Melbourne, Florida, USA, March 09-12, 2003, ACM New York (2003), 760–764.

[6] T. Wang, Y. Rui, J.-G. Sun, Constraint Based Region Matching for Image Retrieval, *Int. J. Comput. Vision* **56** (2004), no. 1-2, 37–45.

[7] M. Mutica, A. Ciubara, L. Duica, D. Alexandru, C. Plesea, M. Pirlog, M. Cara, Elderly schizophrenic patients - clinical and social correlations, *European Neuropsychopharmacology* **26** (2016), no. 2, 5512–5512.

[8] M.M. Trivedi, C.A. Harlow, R.W. Conners, S. Goh, Object detection based on gray level cooccurrence, *Comput. Vision Graph. Image Process.* **28** (1984), no. 3, 199–219.

[9] S. Singh, M. Markou, An Approach to Novelty Detection Applied to the Classification of Image Regions, *IEEE Trans. on Knowl. and Data Eng.* **16** (2004), no. 4, 396–407.

[10] N, Bassiou, C. Kotropoulos, Color image histogram equalization by absolute discounting back-off, *Computer Vision and Image Understanding* **107** (2007), no. 1-2, 108–122.

[11] C. Silisteanu, L. Mitariu, R. Ranga, E. Antonescu, L. Duica, M. Racheriu, M. Totan, M. Manea, Potentiating the Effect of Treatment with Voltaren Gel Using Ultrasonic Frequencies of 1 MHz, *Revista de Chimie* **69** (2018), no. 7, 1749–1751.

[12] M. Totan, E. Antonescu Elisabeta, G. Catana, M. Mitariu, L. Duica, C. Filip, R. Comaneanu, C. Mitariu, C-Reactive Protein - A Predictable Biomarker in Ischemic Stroke, *Revista de Chimie* **70** (2019), no. 6, 2290–2293.

[13] K. Onoguchi, Moving Object Detection Using a Cross Correlation between a Short Accumulated Histogram and a Long Accumulated Histogram, *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, IEEE Computer Society, Washington, DC, USA, 2006, 896–899.

[14] Z. Cao, F. Chen, Y. Du, Object Detection Using Geometrical Block Structures, *ICNC '07: Proceedings of the Third International Conference on Natural Computation* (ICNC 2007), IEEE Computer Society, Washington, DC, USA, 2007, 561–565.

(Nicolae Constantinescu) Department of Computer Science, University of Craiova, 13 A.I. Cuza Street, Craiova, 200585, Romania
E-mail address: nikyc@inf.ucv.ro