Image Processing Using Pearson’s Correlation Coefficient: Applications on Autonomous Robotics
Arthur Miranda Neto, Alessandro Corrêa Victorino, Isabelle Fantoni, Douglas Eduardo Zampieri, Janito Vaqueiro Ferreira, Danilo Alves Lima

To cite this version:

HAL Id: hal-00860912
https://hal.archives-ouvertes.fr/hal-00860912
Submitted on 11 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract—Autonomous robots have motivated researchers from different groups due to the challenge that it represents. Many applications for control of autonomous platforms are being developed and one important aspect is the excess of information, frequently redundant, that imposes a great computational cost in data processing. Taking into account the temporal coherence between consecutive frames, we have proposed a set of tools based on Pearson’s Correlation Coefficient (PCC): (i) a Discarding Criteria methodology was proposed and applied as (ii) a Dynamic Power Management solution; (iii) an environment observer method based on PCC selects automatically only the Regions-Of-Interest; and taking place in the obstacle avoidance context, (iv) a method for Collision Risk Estimation was proposed for vehicles in dynamic and unknown environments. Applying the PCC to these tasks has not been done yet, making the concepts unique. All these solutions have been evaluated from real data obtained by experimental vehicles.

I. INTRODUCTION

In the last three decades, visual navigation for mobile robots or unmanned vehicles has become a source of countless research contributions. Some of these applications include: the development of Unmanned Aerial Vehicles (UAVs) that has been of interest for military applications. However, one limitation is their maximum flight time: they cannot carry large fuel payloads [3]. Future exploration of Mars requires long-endurance UAVs that use resources that are plentiful on Mars [4]. Finally, for military or civil purposes, vehicular applications have as objective the development of autonomous systems capable of driving a car autonomously in an urban environment and also to help the driver in its driver task [5], [6]. All these real-time platforms must provide the capability of perceiving and interacting with its neighbour environment, managing power consumption, CPU usage, etc.

The primary interest in this work, which is environment perception, is still in evolution. Part of this, machine vision, is an important tool that continuously monitors the way forward, proving appropriate road information in real time. Although extremely complex and highly demanding, thanks to the great deal of information it can deliver, machine vision is a powerful means for sensing the environment and it has been widely employed to deal with a large number of tasks in the automotive field [7], but it can lead to some losses due to the processing time.

The problems of time-dependent and dynamic resource allocation have manifested themselves under different names, which include energy and memory consumption for the embedded systems [8]. It has been a topic of interest in the automotive industry [9].

In 1885, an empirical and theoretical development that defined regression and correlation as statistical topics were presented by Sir Francis Galton [10]. In 1895, Karl Pearson published the Pearson’s Correlation Coefficient (PCC) [11]. The Pearson's method is widely used in statistical analysis, pattern recognition and image processing [12].

Based on Pearson's method, we have proposed a visual-perception system based on an automatic image discarding method as a simple solution to improve the performance of a real-time navigation system by exploiting the temporal coherence between consecutive frames [13]. This idea is also presented in the key-frame selection technique [14]. Further, we present the PCC as an environment observer method to save processor energy (power) consumption [15]. In the obstacle avoidance context for vehicles in dynamic and unknown environments, we have also proposed two others methods: a real-time collision risk estimation [16] and an extension of the environment observer method that selects automatically only the Regions-Of-Interest (ROI) [17].

To better understand these toolkit proposals, the Pearson's method is presented in Section II, followed by the Discarding Criteria method in Section III. Section IV introduces the Visual-Perception Layer based on Monocular Vision. Thereafter, the following sections are: Section V: Real-Time Dynamic Power Management, Section VI: Automatic Regions-Of-Interest Selection and Section VII: Collision Risk Estimation. Finally, the results are presented in Section VIII and the conclusions are given in Section IX.

II. PEARSON’S CORRELATION COEFFICIENT (PCC)

The Pearson's method is widely used in statistical analysis, pattern recognition and image processing. Applications on the latter include comparing two images for image registration
purposes, disparity measurement, etc [12]. It is described in (1):

$$r_i = \frac{\sum (x_i - x_m)(y_i - y_m)}{\sqrt{\sum (x_i - x_m)^2 \sum (y_i - y_m)^2}}$$

where \(x_i\) is the intensity of the \(i^{th}\) pixel in image 1, \(y_i\) is the intensity of the \(i^{th}\) pixel in image 2, \(x_m\) is the mean intensity of image 1, and \(y_m\) is the mean intensity of image 2.

III. DISCARDING CRITERIA

The discarding criteria was presented as a simple solution to improve the performance of a real-time navigation system by choosing, in an automatic way, which images should be discarded and which ones should be treated at the visual perception system [13]. It was a new approach using the PCC.

In Fig. 1, basically, if the PCC indicates that there is a high correlation between a reference frame and another new frame acquired, the new frame is discarded without being processed (for example, the system can repeat a last valid command). Otherwise, the frame is processed and it is set as the new reference frame for the subsequent frame.

The inclusion of an automatic image discarding method leads in a reduction of the processing time. Although the system spends some milliseconds computing the PCC, it gains much more time, in some cases, discarding more than 90% of the images [18]. However, it is important to notice that this percentage is not dependent on the video sequence or image size, but on the obstacles / objects influence.

![Fig. 1 – Discarding criteria [13.]](image)

IV. VISUAL-PERCEPTION LAYER BASED ON MONOCULAR VISION

The perception of the environment is a major issue in autonomous vehicles. The perception layer uses many types of sensors [5]. The vision-based sensors are defined as passive sensors and the image scanning is performed fast enough for mobiles robots. However, vision sensors are less robust than millimeter-wave radars in foggy, night, or direct sun-shine conditions [7]. All range-based obstacle detection systems have difficulty for detecting small or flat objects on the ground, and range sensors are also unable to distinguish between different types of ground surfaces [19]. However, the main problem with the use of active sensors is represented by interference among sensors of the same type, hence, foreseeing a massive and widespread use of these sensing agents, the use of passive sensors obtains key advantages [7].

On the safety front, the progressive safety systems will be developed through the manufacturing of an “intelligent bumper” peripheral to the vehicle in answering new features as: blind spot detection, frontal and lateral pre-crash, etc. The objective in terms of cost to fill ADAS functions has to be very lower than the current Adaptive Cruise Control (500 €) [20].

Aware that in the majority of the autonomous systems, the machine-vision system is working together with other sensors, added to its low cost, this work uses a monocular vision-based sensor. Because it uses simple techniques and fast algorithms, the system is capable to achieve a good performance, where the compromise between processing time and images acquisition is fundamental.

V. REAL-TIME DYNAMIC POWER MANAGEMENT

A. Related Work

Autonomous robots can perform desired tasks in unstructured environments without continuous human guidance. These systems have some degree of self-sufficiency. Self-configuring, self-optimizing and self-protecting are still open questions. For advances in the energy autonomy, robots will need to extract energy from the environment. In many ways robots will face the same problems as animals [9].

In this way, a system must therefore have knowledge of its available resources as well as its components, their desired performance characteristics and their current status. Dynamic Power Management (DPM) is a design methodology for dynamically reconfiguring systems to provide the requested services and performance levels with a minimum number of active components or a minimum load on such components. It encompasses a set of techniques that achieves energy-efficient computation by selectively turning off (or reducing the performance of) system components when they are idle (or partially unexploited) [21]. An autonomous robot planning tasks must be aware of power resources available [9]. Moreover, most electronic circuits and system designs are confronted with the problem of delivering high performance with a limited consumption of electric power, and for achieving highly energy-efficient computation is a major challenge in electronic design [21]. In this context, a DPM and Real-Time Scheduling techniques were presented in [22]. They were applied to reduce the power consumption of mobile robots. The DPM dynamically adjusts power states of components adaptive to the task’s need, reducing the power consumption without compromising system performance.

Finally, a case study of mobile robot’s energy consumption and conservation showed that motion accounts for less than 50% of the total power consumption [22]. This implies that other power consumers like computation have a big impact on power consumption [9]. These values were estimated by dividing the battery capacity by the time the computer can run with a fully charged battery when running different programs [22].
B. Logical Dynamic Optimization

This section presents a logical dynamic optimization methodology. Based on the PCC variation and by exploiting the temporal coherence between consecutive frames, it is proposed a new environment observer method. This monocular-vision system observes if there are no significant changes in the environment, permitting that some logical components may be shut down to save processor energy consumption, and/or to make the CPU available for running concurrent processes.

![Fig. 2](image)

**Fig. 2** – (a): The frames of the desert video [23]; (b) From a reference frame, its correlation with all others; Blue line: the Pearson’s correlation in (1); The vertical black line: maximum point before collision; The vertical red line: Empirical Risk-of-Collision.

A robot can have many periodic tasks, such as motor and sensor control, sensing data reading, motion planning, and data processing. It may also have some aperiodic tasks, such as obstacle avoidance and communication. Moreover, for mobile robots, the tasks’ deadlines are different at different traveling speeds. At a higher speed, the periodic tasks have shorter periods [22].

The Fig. 2 (a) shows an autonomous displacement through the Mojave Desert [23], where the robot Stanley has used an average speed of 30.7 km/h [24]. In Fig. 2 (b), due to PCC nature, taking a reference frame, in this case, the first frame of the Fig. 2 (a), a lower value of correlation is achieved when it is closer to the vehicle. That is, when the derivative approaches its maximum point, there is the obstacle detection.

The Fig. 3 (a) shows the same case from a different representation. From an empirical PCC threshold equal to 0.85 (green line), the reference frames (red points) are closer when it is near to an obstacle. Above of the green line all discarded images.

Whereas the main problem that has to be faced when real-time imaging is concerned and which is intrinsic to the processing of images is the large amount of data [7], and as was presented in [15], the accumulated time of an image processing time (15ms) versus the gain obtained by using the discarding criteria could allow significant savings in CPU power consumption. In this case, the discarding rate remained over 80%.

VI. AUTOMATIC REGIONS-OF-INTEREST SELECTION BASED ON PCC (ROI SELECTION)

According to the Pearson’s correlation, in a certain analysis window (pair of frames), if the obstacle/object occupies a big portion of the scene, the PCC threshold tends to be low. Conversely, if obstacle/object occupies a small portion of the frame, it means that it is away from the vehicle and the system will have time enough to react. However, in real-time obstacle avoidance, for example, where are these interest points/pixels? Or, in a sequence analyzed, which pixels of the pair of images contributed most to the Pearson’s coefficient computed? Which of them really need to be reprocessed?

Right after the Pearson’s correlation in (1), it has $x_m$ and $y_m$, respectively: the mean intensities of images 1 and 2. From these values, it begins again the process’s correlation:

$$ r_2 = \frac{\sum (x_i - r_{1xm})(y_i - r_{1ym})}{\sqrt{\sum (x_i - r_{1xm})^2} \cdot \sqrt{\sum (y_i - r_{1ym})^2}} = \begin{cases} -1 & \text{or} \\ +1 & \end{cases} $$

where $x_i$ is the intensity of the $i^{th}$ pixel in image 1, $y_i$ is the intensity of the $i^{th}$ pixel in image 2, $r_{1xm}$ and $r_{1ym}$ were obtained in (1): i.e.: $x_m$ and $y_m$.

For each pair of pixels analyzed in (2), the only possible result is: [-1 or +1]. That is, all pixels with intensities below these means will be candidates for interest points (ROI). Fig. 4 (c), (g) and (k) present this process, where the red pixels (interest points) represent $r_2 = -1$.

Taking as base an image resolution equal to image 96x72, by processing only when $r_2 = -1$, in desert video were processed about 205 thousand points, instead of 3.7 million points. In off-road context were processed about 10 million points, instead of 48 million points [17], [25].

VII. COLLISION RISK ESTIMATION

A. Related work

In the obstacle avoidance context, the collision warning algorithms typically issue a warning when the current range to an object is less than the critical warning distance, where the safety can be measured in terms of the minimum time-to-collision (TTC) [26]. To calculate the TTC several techniques...
are presented in the literature [27], [28]. Measuring distances is a non-native task for a monocular camera system [27]. However, TTC estimation is an approach to visual collision detection from an image sequence.

Optical flow may be used to TTC [29], [30]. However, the calculation of TTC from an optical flow has proven impractical for real applications in dynamic environment [28]. Additionally, gradient-based methods can be used with a certain degree of confidence in environments such as indoors where the lighting conditions can be controlled. It is computationally expensive [31]. On the other hand, Inverse Perspective Mapping allows to transforms a front facing image to a top down bird’s eye view [32]. However, those equations have parameters that depend on the camera’s position and its viewing angle [33].

In this way, we have presented a novel approach to obtain Collision Risk Estimation (CRE) based on PCC from a monocular camera [16].

![Fig. 4 – (a), (b), (e), (f), (i) and (j) are the frames of the desert video [23]; (c), (g) and (k) are the interest points from the process’s correlation in (2); (d), (h) and (l) are the ITA results [17].](image)

**B. Collision Risk Estimation (CRE)**

By exploiting the temporal coherence between consecutive frames, this section presents an algorithm which estimates the CRE in dynamic and unknown environments.

The Fig. 4 (a) shows an autonomous displacement through the Mojave Desert [23], where the robot Stanley has used an average speed of 30.7 km/h [24]. In Fig. 2 (b), taking a reference frame, i.e. the first frame of the Fig. 4 (a), a lower value of correlation is achieved when it is closer to the vehicle, Fig. 2 (b): black line. That is, when the derivative approaches its maximum point, there is the obstacle detection. Fig. 2 (b): red line presents an Empirical Risk-of-Collision, \( R = 1 - 0.6 \).

Taking into account \( R \), the CRE is estimate in (3):

\[
CRE = \frac{R}{(1 - r_1)}
\]

where \( 1 \) (one) represents the reference frame and \( r_1 \) was obtained in (1) and \( R = 1 - 0.6 \).

**C. Obstacle Direction: Interactive Thresholding Algorithm**

From the interest points known in Section VI, this section presents the Interactive Thresholding Algorithm (ITA) [17] that reclassifies the background and foreground pixels based on Otsu Thresholding Method [34].

Taking as base an image resolution equal to image 96x72, the ITA process will be performed N-times until the result is invariably, or until the red points (foreground) are less than 100. For example, Fig. 4 (d), (h) and (l) present the final result of this process, where the green points were eliminated in the last iteration. Finally, the blue line indicates the object direction based on the center of area of the red points.

**D. Collision Risk Estimation: Case Study**

The Table I presents the CRE from the Fig. 4:

- Frames column: it represents the pairs of frames [1-30], [300-330] and [400-430], respectively: Fig. 4: [(a), (b)], [(e), (f)] and [(i), (j)].
- \( (1- r_1) \) column: the Pearson’s correlation obtained in (1).
- Variation in the Range column: the PCC variation between the first and last frames analyzed.
- CRE Second column: it estimates in (3).
- Distance Meters column: it presents an estimate in meters from the average speed of 30.7 Km/h [24].

**TABLE I**

<table>
<thead>
<tr>
<th>Frames</th>
<th>((1- r_1))</th>
<th>Variation in the Range</th>
<th>Risk of Collision</th>
<th>CRE in Seconds</th>
<th>Distance in Meters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-30</td>
<td>0.032</td>
<td>((R / 0.032))</td>
<td>12.43s</td>
<td>106m</td>
<td></td>
</tr>
<tr>
<td>300-330</td>
<td>0.093</td>
<td>((R / 0.093))</td>
<td>4.31s</td>
<td>36.75m</td>
<td></td>
</tr>
<tr>
<td>400-430</td>
<td>0.200</td>
<td>((R / 0.200))</td>
<td>2.00s</td>
<td>17.08m</td>
<td></td>
</tr>
</tbody>
</table>

**VIII. EXPERIMENTAL RESULTS**

Besides the experimental DARPA test-banks [23], the results here were obtained using an experimental vehicle (Fig. 19) on real, dynamic and unknown environments. It was tested on a 2.5GHz Intel Core 2 Quad processor, 3.48 GB RAM, Microsoft Windows XP Professional SP3, Visual Studio C++ and OpenCV 2.1.0. In order to reduce the number of data, it also includes the resolution reduction of image (to 96x72).

Following the same structure presented earlier:

- Subsection A: Real-Time Dynamic Power Management;
- Subsection B: Automatic Regions-Of-Interest Selection;
- Subsection C: Collision Risk Estimation.
A. Section V: Real-Time Dynamic Power Management

The Figures 15 and 16 show the performance of this method in real, dynamic and unknown environments. For all these cases, the discarding rate remains over 65%. Fig. 16 (a) presents the computational mean time of a horizon finding algorithm [35] in unknown and urban environment. In this way, from an empirical PCC threshold equal to 0.85, the red line shows that the computational mean time was 5.09 ms, against 15.62 ms without the discarding criteria. In Fig. 17 (b), above the green line, it presents the discarded images. A video showing the application of this method is available in [36].

B. Section VI: Automatic Regions-of-Interest Selection

As has been shown in Section VI, at first stage of testing, in order to evaluate the proposed algorithm performance, it was used an urban and real experimental test-bank obtained using the vehicle shown in Fig. 19. Results for different types of image texture (road surfaces) were selected and its results are presented in [25]. For obstacle avoidance task, the Fig. 17 presents results at high speed on real-time conditions. A video showing the application of this method is available in [37].

C. Section VII: Collision Risk Estimation

Fig. 18 and Table II present the performance of the Collision Risk Estimation (CRE) in dynamic and unknown environment. These results were obtained in real conditions using the vehicle shown in Fig. 19.

Since in real conditions this monocular-vision system has been designed to investigate only a small portion of the road ahead of the vehicle, where the absence of other vehicles has been assumed [11], the Fig 18.a-*(a) presents the fix analysis region (yellow line). As shown in [16], the computational mean time of the CRE process was equal to 7.8 ms. A video showing the application of this method is available in [38].

<table>
<thead>
<tr>
<th>Frames</th>
<th>Variation in the Range</th>
<th>Risk of Collision</th>
<th>CRE in Seconds</th>
<th>Distance in Meters</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>(1-0.8315)</td>
<td>0.1685</td>
<td>(R / 0.1685)</td>
<td>2.37s</td>
</tr>
<tr>
<td>(b)</td>
<td>(1-0.5584)</td>
<td>0.4416</td>
<td>(R / 0.4416)</td>
<td>0.90s</td>
</tr>
<tr>
<td>(c)</td>
<td>(1-0.5411)</td>
<td>0.4589</td>
<td>(R / 0.4589)</td>
<td>0.87s</td>
</tr>
<tr>
<td>(d)</td>
<td>(1-0.7394)</td>
<td>0.2606</td>
<td>(R / 0.2606)</td>
<td>1.53s</td>
</tr>
</tbody>
</table>

IX. Conclusion

This work presents a simple solution to improve the performance of a real-time perception system. The experiments showed that the inclusion of an automatic image discarding method based on PCC did result in a reduction of the processing time. This technique is also presented as an environment observer method (DPM) and futures work will...
provide a real experimental test-bank to evaluate the real energy consumption economy in terms of electrical current used by the visual machine. Following, instead of processing all image pixels, an extension of the DPM that selects automatically only the ROI was proposed in order to perform an obstacle avoidance task in real time. Finally, a real-time algorithm which allows to calculate the risk of collision. In order to validate this proposal, futures work would be focused to provide ground truth measurements from a front mounted radar and/or LIDAR system. A remarkable characteristic of all methodologies presented here is its independence of the image acquiring system and of the robot itself. The same implementation can be uses in different mobile robots and may be extended to other sensors.

Fig. 19 – The experimental vehicle at Heudiasyc Laboratory in France.

ACKNOWLEDGMENT
The authors wish to thank Mr. Gerald Dherbomez for your support in data acquisition.

REFERENCES
[38] [http://you.be/J8YuZjUFEsk [Jan. 31, 2013]