Face Detection on Infrared Thermal Image

Ricardo F. Ribeiro, José Maria Fernandes, António J. R. Neves

DETI/IEETA, University of Aveiro,
3810-193 Aveiro, Portugal
Email: {rfribeiro,jfernan,an}@ua.pt

***Abstract*—Infrared cameras or thermal imaging cameras are devices that use infrared radiation to capture an image. This kind of sensors are being developed for almost a century now. They started to be used in the military environment, but at that time it took too long to create a single image. Nowadays, the infrared sensors have reached a whole new technological level and are used for purposes other than military ones, as happens in this work, where they are being used for face detection. When comparing the use of thermal images regarding color images, it is possible to see some advantages and some limitations, which will be explored in this paper. This work proposes the development or adaptation of several methods for face detection on infrared thermal images. The well known algorithm developed by Paul Viola and Michael Jones, using Haar feature-based cascade classifiers, is used to compare the traditional algorithms developed for visible light images when applied to thermal imaging. In this paper, we present three different methods for face detection. As far as we know, there is limited research on this topic so we think this work is an important contribution to the field. In the first one, an edge detection algorithm is applied to the binary image and the face detection is based on these contours. In the second method, a template matching method is used for searching and finding the location of a template image with the shape of human head in the binary image. In the last one, a matching algorithm is used. This algorithm correlates a template with the distance transform of the edge image. This algorithm incorporates edge orientation information resulting in the reduction of false detection and the cost variation is limited. The results show that the proposed methods have promising outcome, but the second method is the most suitable for the performed experiments.**

***Keywords*–Face detection; infrared images;Image processing; robotics; object detection.**

I. Introduction

In the electromagnetic spectrum, the visible light spectrum is the only part that the human eye can see. Due to the fact that infrared radiation is invisible to the human eye, thermal cameras use infrared sensors to capture that radiation, transforming it into visible images. Many objects and even humans emit infrared radiation in function of the temperature: the higher the temperature, the higher the intensity of the emitted radiation.

The use of thermal infrared cameras has been increasing in various scientific areas. A survey providing an overview of the current applications is presented in [1]. Applications include animals, agriculture, buildings, gas detection, industrial, and military fields, as well as detection, tracking, and recognition of humans. In robotics, for computer vision, thermal image analysis and processing is in constant development, being used more often in systems for detection and tracking of objects, humans, among others.

The camera used in this work, which is shown in Figure 1, is a complete long-wave infrared (LWIR) camera module that captures infrared radiation input in the wavelength range from 8 to 14 microns and converts it to infrared thermal image. Figure 2 shows a thermal image complemented with an image in the visual spectrum from the same scene (typically a Red, Green and Blue - RGB - image).

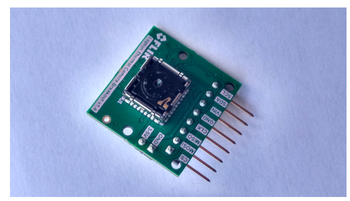

Figure 1. FLIR Lepton thermal camera module with breakout board.

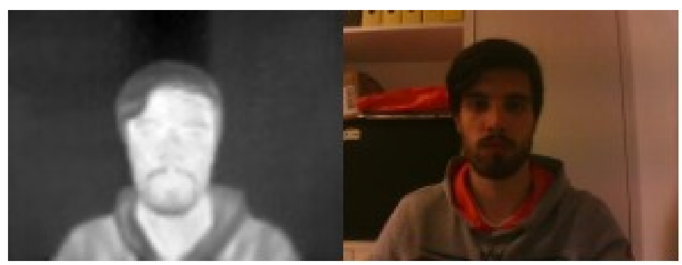

Figure 2. An example of a thermal image and the same scene aquired by a RGB camera.

In this work, we propose algorithms for face detection using thermal infrared cameras. The main goals of this work are the use of these type of sensors in service robots and in monitoring people attention taking into consideration of the temperature of the face over time. We are also working on emotional analysis through thermal images.

This paper is organized as follows. In Section II, we describe some advantages and limitations of thermal images. In Section III, we present the techniques that have been proposed for face detection. In Section IV, we provide experimental results. Finally, in Section V, we draw some conclusions.

II. Thermal Images

Thermal cameras are advantageous in many applications due to their ability to see in total darkness, their robustness to illumination changes and shadow effects, and less intrusion on privacy [2]. These cameras, when calibrated, are advantageous in temperature measurement compared to point-based methods

 ISBN: 978-1-61208-559-3

since temperatures over a large area can be compared, although contact methods are more efficient [3]. In this work, the thermal camera used is not calibrated, so it is not possible to know the exact temperature in the region of interest.

Some algorithms for face detection using color or grayscale images currently have a very high efficiency rate but is not possible to use them directly in the thermal images. In thermal imaging there are some problems that can reduce this efficiency significantly using these algorithms such as occlusion of the face with objects that emit infrared radiation, the uniform temperature in the face, objects with the shape and same temperature of the face, among others. For this reason, only one feature-based algorithm is used, the Haar Cascade [4].

Consequently, with problems regarding face detection, the temperature becomes unstable over time. The face is one of the zones of the human body that is more suitable for the body temperature extraction and posteriorly the emotion analysis. The face detection in this work is represented on a shape of a bounding box obtaining the face and also some surrounding background noise that can affect the detected face for the next modules. The temperature measurement and emotional analysis of the face is being developed, not being included in this current work.

III. Proposed Approach

In this section, different developed algorithms and methods for face detection on thermal images are described. The development of this work was made in C++ programming language, through the use of the OpenCV library, which is an open source computer vision and machine learning software library. The use of these algorithms and methods will be on real-time systems, therefore it is also described in this section the equipment used for the acquisition of the thermal image and the method to obtain these images.

A. Image Acquisition

For thermal image acquisition in real-time, it is used a FLIR LEPTON Long Wave Infrared (50 shutterless) camera module, with a focal plane array of 80x60 active pixels. This camera is a non-radiometric version. This camera is also not calibrated, therefore it is not possible to obtain temperature values of each pixel. The output value also changes with the temperature value of the infrared sensor of the camera and the temperature of the scene.

It is also used a Raspberry Pi 3 model B for communication and image processing. For the acquisition of the output values on the Raspberry Pi 3, Serial Peripheral Interface (SPI) communication is used. It also supports a command and control interface (CCI) hosted on a Two-Wire Interface (TWI) similar to Inter-Integrated Circuit (I2C) for software interface [5]. This image acquisition process is based on a project developed by the company Pure Engineering [6]. The obtained output values are received in an 14-bits data, then they are arranged in an 8-bits with one channel image matrix format provided by OpenCV. The image on the left presented in Figure 2 shows an example of a thermal image. Darker areas correspond to colder regions in the scene.

B. Haar Cascades

Haar Cascades is a machine learning algorithm where a cascade function is trained from positive and negative images. This approach uses the Viola and Jones algorithm [4].

Haar Cascades is one of the algorithms implemented by OpenCV library. This algorithm was studied by Mekyska et al. [7] showing a machine learning approach for face detection and it requires a high number of images of the object to be detected for the cascade training. A similar study was made in this work, but the results of the face detection are considered of low accuracy. The thermal images with faces that were used for the cascade training can be obtained on the online dataset [8]. The result of the cascade training for Haar Cascade is shown in the results section.

C. Implementation Details

As far as we know, face detection on thermal images did not received too much attention on computer vision as the counterpart on visible light images. There are some possible ways using some functionalities of OpenCV library. Thermal image is, on a first stage, segmented and filtered with morphological operators in order to obtain a binary image for the later use of some methods or algorithms.

Segmentation uses the Otsus method that is a thresholding binarization method [9] and filtering is performed using morphological operators, such as dilation, erosion, opening and closing [10]. We developed and implemented the following algorithms:

- **Face Contours** - Acquisition and filtering of the contours in order to obtain the longest contour in the binary image and detect the face through it [11].
- **Template Matching** - Technique for finding areas of an image that match to a template image [12].
- **Chamfer Matching** - Technique to find the best alignment between two edge maps [13].

1) Face Contours: Through the binary image is created an edges map, using the Canny edge detector algorithm [14], where the contours are found [11]. These contours are filtered in order that only the contour of larger area are obtained. Due to the contour of having some parts of the human body that are not relevant for face detection, for example neck and shoulders, it is found the highest point of the contour. This point matches to the highest point on the face. Starting at this point, the two points that correspond to the largest width of the face are found. The detection of the face is made with this two points and the highest point in the face contour.

2) Template Matching: Template Matching is a technique that uses a template to search and find in an image the best match of this template. There are different matching methods to perform the template matching technique. This work uses the Normalized Cross-Correlation method that remains a viable choice for some if not all applications [12].

Researchers in [15] use a Image Pyramid to perform the template matching for objects with different sizes. In this work it is used image pyramid to detect faces of different sizes increasing the performance of the template matching. Due to the image captured by the camera having a reduced resolution, few levels of the pyramid image are used. Each level of the image is downsized. A good template is needed to obtain better results in the template matching.

 ISBN: 978-1-61208-559-3

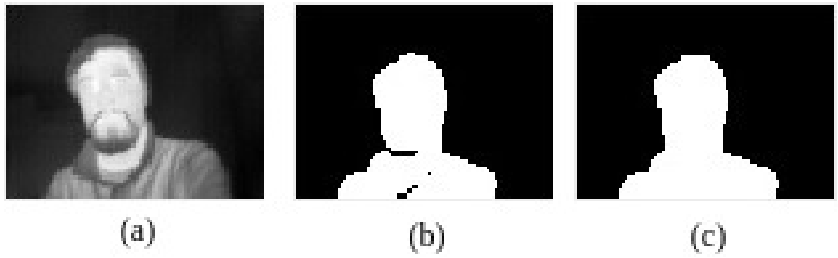

Figure 3. In (b) the segmentation of the thermal image (a). The result of applying the morphological operation is presented in (c).

3) Chamfer Matching: The Chamfer Matching algorithm lies on shape matching using distance transform. This technique uses an edges image of the image being tested and the template to create an Distance Transform map [16]. The value of each pixel in this map is the distance to the nearest background pixel. After this, chamfer distance map is created that contains the computed matching cost through the distance transform maps and the position of the edge orientation contours [13]. The face is detected finding the pixel location of the minimal chamfer distance.

IV. Results

In this section, we present experimental results to verify the effectiveness of the developed algorithms. Since our goal is the real time use of this contribution, the examples were obtained in real time with the camera connected to the single board computer. The obtained output values used to form the thermal images were processed in order to obtain an 8-bit grayscale image. An example of this type of images is represented in Figure 3(a).

A. Haar Cascades

The study made by Mekyska et al. [7] used the Viola and Jones algorithm [4] applied to thermal image. The results of this study show that in order to obtain a good accuracy, a large amount of training data is needed. The disadvantage of this algorithm is also the detection time, which is dependent on the amount of training data.

In this work, there was an attempt of performing cascade training for face detection. An example of the application of this algorithm is presented in Figure 4. However, in comparison with the study mentioned it has an inferior performance. Face detection is unstable since it does not detect the face if the image does not contain the neck and shoulders of the person.

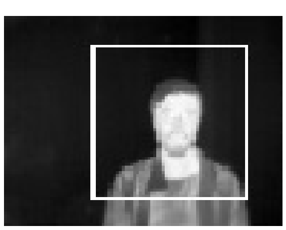

Figure 4. Face detected using Haar Cascades.

B. Proposed Methods for Face Detection

Figures 3(b) and 3(c) show and example of the segmentation obtained in this type of images and the use of morphological operators, respectively.

The methods and algorithms mentioned in Section 3.3 use and input the image presented on Figure 3(c).

1) Face Contours: Using Canny algorithm the edge map presented in Figure 5 is obtained. Figure 6 shows an example of face detection using contours. These results can be influenced if the contours of the face are discontinued for some reason. This method is not the most suitable for face detection since there are some variations in the bounding box of the detected face.

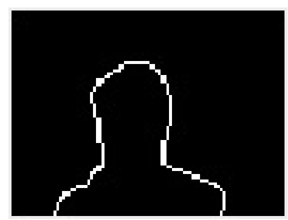

Figure 5. Edge map obtained from a thermal image.

2) Template Matching: When using template matching a good template is needed. Figure 9 shows the template used. This method slides the template over the image of Figure 3(c) and calculates an error for the match between the template and the image being tested [15]. A method based on image pyramids is also used in order to improve the result of template matching since it takes the scale in consideration. The best match is found as a global maximum value. Figure 7(b) shows some examples of application.

3) Chamfer Matching: This algorithm uses the same template of Figure 9, but it is converted to an edge using the Canny edge detector algorithm to be used later in the construction of the cost image. A distance transform map is created from the edge image and from the template, which specifies the distance from each pixel to the nearest edge pixel in the query image.

Figure 10 represents the two images of distance transform maps. Then a map with the matching cost of each pixel is created. The pixel location of the minimal cost is the location of the region of interest. Some results applying the Chamfer Matching [13] are shown in Figure 8.

C. Processing Time

All experimental results have been obtained in real time using a Raspberry Pi 3 model B.

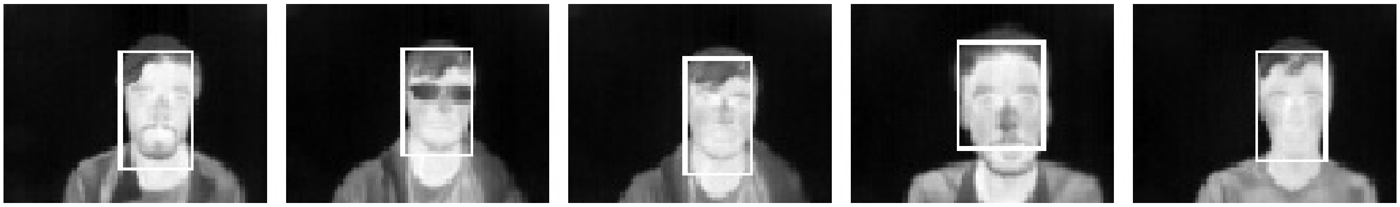

Figure 6. Some examples of face detection in different conditions using contour detection.

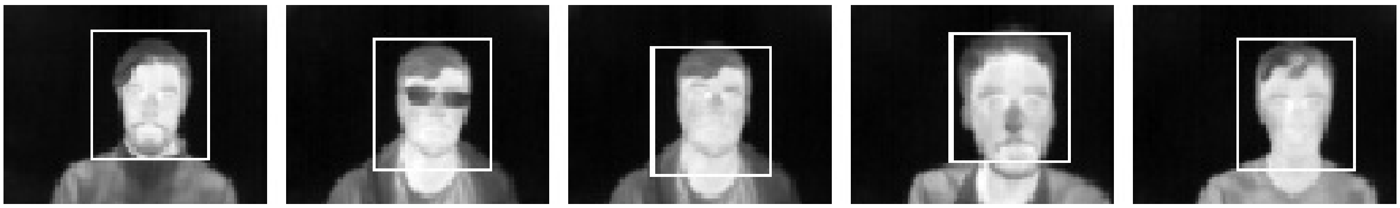

Figure 7. Some examples of face detection in different conditions using Template Matching.

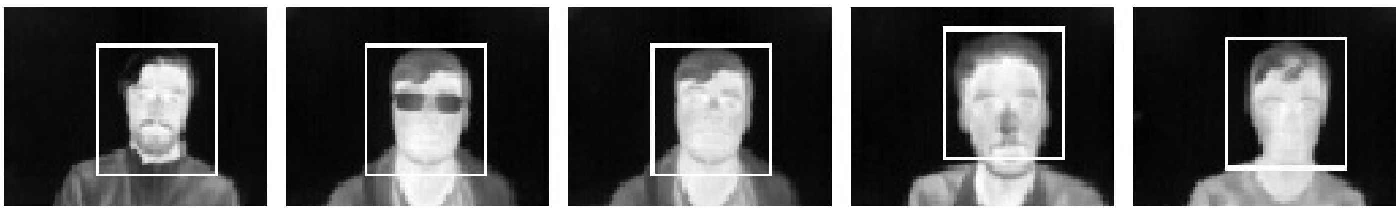

Figure 8. Some examples of face detection in different conditions using Chamfer Matching.

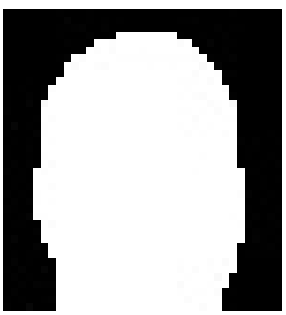

Figure 9. Template used for template matching.

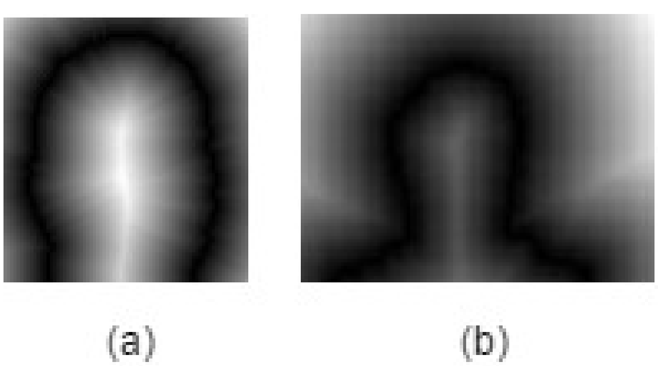

Figure 10. Distance Transform image of template (a) and query image (b).

The processing time of Haar Cascade is approximately 72ms.

The average processing time obtained by the proposed methods applied during an experiment of 70 seconds with a frame rate limited to 4 frames per seconds are the following:

- **Face Contours** - 68ms.
- **Template Matching** - 69ms.
- **Chamfer Matching** - 257ms.

Chamfer Matching has the highest processing time, because this method involves several steps which require more processing time, for example distance transform, edge orientation, among others.

V. Conclusion and Future Work

In this paper, we presented a study regarding the development of three different methods for face detection in thermal images using image segmentation.

Haar Cascade using Viola and Jones algorithm has better performance and accuracy based on [7]. However, this algorithm needs a large amount of data and time for training to obtain good results. Face detection in thermal image using Haar Cascade can be improved using one of the proposed methods for image segmentation and create a lot of thermal images that contain only the face for training database.

Face detection through thermal imaging using segmentation has a good accuracy in single face detection. Comparing the proposed methods, Template Matching is the most suitable. Although Face contours has a processing time similar to Template Matching, the bounding box with the face detected is more unstable. Chamfer Matching has a similar detection to Template Matching but processing time of this method is almost 4 times slower.

In this work, the thermal camera used is not calibrated, so it is not possible to know the exact temperature in the region of interest. As future work, we intend to improve the developed algorithms for face detection and develop algorithms for calibration of this sensor in order to measure absolute temperatures. We are also working on a calibration procedure for the simultaneous use of this camera and an RGB camera.

REFERENCES

[1] R. Gade, and T. B. Moeslund, ”Thermal cameras and applications: a survey”, *Machine vision and applications* 25.1: 245-262, 2014.

[2] A. Berg, J. Ahlberg and M. Felsberg, ”A thermal object tracking benchmark”, *Advanced Video and Signal Based Surveillance (AVSS), 12th IEEE International Conference on.* IEEE, 2015.

[3] A. Berg, *Detection and Tracking in Thermal Infrared Imagery* Diss. Linkping University Electronic Press, 2016.

[4] P. Viola and M. Jones, ”Rapid object detection using a boosted cascade of simple features,” *Computer Vision and Pattern Recognition, CVPR 2001, Proc. of the 2001 IEEE Computer Society Conference on.* vol. 1. IEEE, 2001.

[5] FLIR LEPTON Long Wave Infrared (LWIR) Datasheet. http://www.flir.com/uploadedFiles/OEM/Products/LWIR-Cameras/Lepton/Lepton%20Engineering%20Datasheet%20-%20without%20Radiometry.pdf, retrieved April, 2017.

[6] Pure Engineering LEPTON project. http://www.pureengineering.com/projects/lepton, retrieved April 2017.

[7] J. Mekyska, V. Espinosa-Dur and Marcos Faundez-Zanuy, ”Face segmentation: A comparison between visible and thermal images”, *Security Technology (ICCST 2010), IEEE International Carnahan Conference on*, IEEE, 2010.

[8] Terravic Facial Infrared Database. http://vcipl-okstate.org/pbvs/bench/Data/04/download.html, retrieved April, 2017.

[9] R. F. Moghaddam and M. Cheriet, ”AdOtsu: An adaptive and parameterless generalization of Otsu’s method for document image binarization”, *Pattern Recognition* 45.6: 2419-2431, 2012.

[10] H. Heijmans, ”Connected morphological operators for binary images,” *Computer Vision and Image Understanding* 73.1: 99-120, 1999.

[11] S. Suzuki, ”Topological structural analysis of digitized binary images by border following”, *Computer vision, graphics, and image processing* 30.1: 32-46, 1985.

[12] J. P. Lewis, ”Fast normalized cross-correlation”, *Vision interface*, vol. 10. no. 1, 1995.

[13] M. Liu, et al., ”Fast directional chamfer matching,”, *Computer Vision and Pattern Recognition (CVPR 2010), IEEE Conference on*, IEEE, 2010.

[14] J. Canny, ”A computational approach to edge detection”, *IEEE Transactions on pattern analysis and machine intelligence* 6: 679-698, 1986.

[15] L. Ferreira, A. J. R. Neves, A. Pereira, E. Pedrosa, J. Cunha, ”Human detection and tracking using a Kinect camera for an autonomous service robot”, Advances in Aritificial Intelligence - Local Proceedings, EPIA 2013 - XVI Portuguese Conference on Artificial Intelligence, Angra do Herosmo, Azores, Portugal, p. 276-288, 2013.

[16] G. Borgefors, ”Distance transformations in digital images”, *Computer vision, graphics, and image processing* 34.3: 344-371, 1986.

 ISBN: 978-1-61208-559-3