Movement Analysis for the Assessment of Hearing in Patients with Cognitive Impairment: A Preliminary Study

A. Fernández-Arias[1], M. Ortega-Hortas[1], B. Cancela-Barizo[1], and L.M. Gigirey[2]

[1] VARPA Group, Department of Computer Science, University of A Coruña, Spain
{alba.fernandez,mortega,brais.cancela}@udc.es
[2] Audiology Unit-University School of Optics and Optometry, University of Santiago de Compostela, Spain
luz.gigirey@usc.es

Abstract. Patients with severe cognitive impairment are not capable of interacting with the audiologist during an audiology test. However, these patients often show gestural reaction to the auditory stimuli that can be interpreted by audiologists with comprehensive experience in this type of case. In this study, we analyze the area around the eyes and the movements that occur within it, since is in this area where most of gestural reactions take place for these patients.

Keywords: Hearing assessment, eyes region detection, movement analysis, optical flow.

1 Introduction

Hearing loss implies a decrease in the sensitivity to sounds that are normally heard, it extends to all age groups with a special emphasis on the elderly. Considering that the World Health Organization (WHO) has set the criteria for hearing loss above 25 dB, it is calculated that more than 55 million European adults suffer from hearing loss in some degree. Future estimates made by Professor A. Davis [1] estimated that by 2015 there will be over 700 million people worldwide with hearing loss over 25 dB, increasing to 900 million people by 2025. Hearing loss and communications disorders may have a negative impact on the state of emotional, physical and social well-being [1] and it also plays a key role in the process of "active aging" [2].

Liminar Tonal Audiometry (LTA) is the gold standard for the clinical evaluation of hearing loss. However, this test may have operational limitations in certain patient groups with "special needs" or cognitive impairment. According to several studies, older adults with hearing loss are more likely to develop Alzheimer's disease and dementia, compared to those with normal hearing. Further, the risk escalates as a person's hearing loss grows worse. The typical interaction during an LTA test is that the patient raises his hand to inform the expert when he

R. Moreno-Díaz et al. (Eds.): EUROCAST 2013, Part II, LNCS 8112, pp. 133–139, 2013.

hears the auditory stimuli. However, when evaluating patients with cognitive decline, the audiologist can not maintain a typical interaction question-answer, so, in most cases, he does not receive voluntary reactions by the patient. This implies that the expert should focus his attention on unconscious facial reactions for what he needs enough experience in order to know how to interpret them. Most of these unconscious facial reactions are shown within the eye region, for example, changes on the gaze direction when perceiving an auditory stimulus, eyes opening as a reaction, frowning or other particular expression changes that could indicate some kind of perception by the patient. It is important to emphasize that each patient may present different expression changes as a reaction, so the audiologist must have broad experience in order to properly explain these gestures. The subjectivity of this task makes of it in a imprecise problem very prone to errors and difficult to reproduce.

In [3] we have previously proposed an automatic system for patients without cognitive impairments in a typical audiological scenario. Therefore, now we need to develop an objective screening method that allow the audiologists to corroborate results if no patient cooperation exits, as in the case of patients with cognitive decline or communication disorders. To that end, at first we want to automatically classify the patient facial reactions. Ir order to achieve this goal, we focus our attention on the area around the eyes because is within this area where most of gestural reactions of these patients take place. This system could be very interesting for the audiologist community, since there are no previous automatic solutions for this task.

By the use of image processing techniques, we detect the eye region and the movements that occur within this region. This initial detection of the movement will serve as a basis for later provide meaning to these movements and be able to interpret them as responses to the auditory stimuli.

This papers is organized into five sections. Section 2 is devoted to briefly describe the critical protocol for the audiometries. Section 3 describes the methodology. Experimental results and some examples are included in Section 4. Finally, section 5 provides some discussion and conclusions.

2 Clinical Protocol for LTA

One of the standard tests for evaluating the hearing capacity is the Liminar Tonal Audiometry (LTA). LTA measures hearing acuity with variations in sound intensity and pitch and for tonal purity. The results of this test determine the least audible sound that a person is able to hear, and they are charted in a graph called audiogram. The inability to hear pure tones below 25dB indicates hearing loss.

The scenario of this test has a very particular setup (as it can be seen in Fig. 1): the patient is located in front of the audiologist, face always in frontal position and wearing earphones connected to an audiometer. With this audiometer the expert delivers pure tones at different frequencies and intensities to the patient's ear, one ear at a time. In a typical interaction, the patient raises his hand to indicate that he has perceived the auditory stimulus. As mentioned

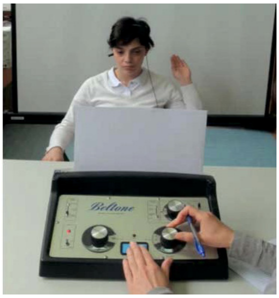

Fig. 1. Sample typical setup during an LTA

during the introduction, in this case we are focusing our attention on patients with cognitive decline, for whom certain gestural reactions are the only response that can to be considered.

In next Section we explain the methodology to evaluate this kind of patients.

3 Methodology

The proposed methodology is a continuation of a previous work [3], where we developed an objective screening system capable of analyzing video sequences recorded during an LTA and to measure the response times of cooperative patients, some of the modules of this system can be reused here. As pointed out by the experts audiologists, most of the facial reactions to the auditory stimuli are concentrated in the eyes area. For this reason, the fist step of this method is to detect this area of interest.

The steps of the proposed methodology are depicted in Fig. 2. Each step is explained in more detail next.

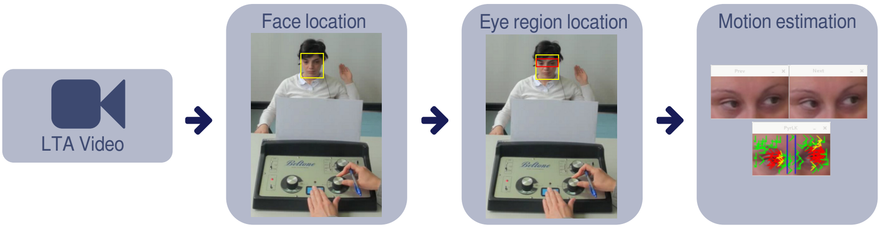

Fig. 2. Schematic representation of the methodology

Face Location. To make faster and less prone to errors the eye region location, we start by locating the face. Since the domain is stable and the particular setup guarantees that faces are always in frontal position, a Viola-Jones [4] approach can be applied. The Viola-Jones object detector framework is a general detection framework that has been trained and optimized for the detection of frontal faces in the OpenCV library. This detector is a low computational solution very robust, therefore, its use is justified.

Eye Region Location. Then, within the detected face area, we need to locate the eye region. To that end, we apply a new Viola and Jones detector specifically trained by us for detecting this specific area. This eye detector was previously trained by us with more than 1000 images of the eye area manually selected (since there was no image database for this region). Tests carried out for the evaluation of this detector showed excellent results (greater than 98% of accuracy) for both open and closed eyes.

Motion Estimation. Once the eye region is detected, now we are interested on detecting and analyzing the movements or expression changes that occur therein. Since we want to see where the movements take place within this region, a point to point registration has no sense, so movement must to be analyzed in a global sense. Moreover, not all the individuals are going to have the same gestures as a reaction to the auditory stimuli and they can also present an erratic behavior. All these particularities make impossible the application of classical solutions based on feature point registration or template analysis. Therefore, it was decided to make a global analysis of the movement without focusing on interest points or facial features.

Our proposal is based on analyzing the optical flow between consecutive eye region images. In this case, we applied the iterative Lucas-Kanade[5] optical flow method with pyramids, which showed good results in the identification of the movements produced by expression changes. Considering that our video sequences have a frame rate of 25 FPS (frames per second) and in order to have expression changes notable enough, we perform comparisons with a tree frame space, meaning, between frame i and frame $i+3$. With a frame rate of 25 FPS, if we compare frame i with frame $i+1$ movements that would occur would be almost imperceptible so they would not be taken in consideration. For that reason, it is better to leave a small gap so we can have movements significant enough to be properly detected.

This calculation of the optical flow is represented on Fig. 3. Figures 3(a), 3(b), 3(c) and 3(d) are four consecutive frames. Leaving a gap of three frames, the optical flow is calculated between 3(a) and 3(d). The results are shown in 3(e) where green vectors represent soft movements, yellow vectors represent intermediate movements and red for strongest movements.

4 Experimental Results

Given the preliminary nature of this study, the aim was to test the viability of using the optical flow for the detection of the facial movements. We must remember that each patient may present different movements as a reaction (even

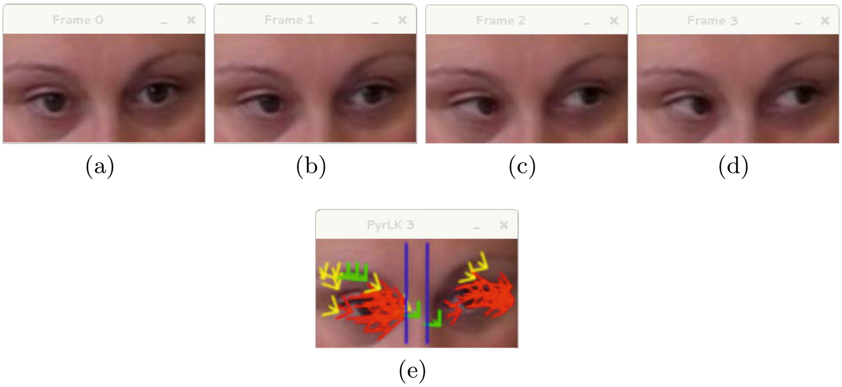

Fig. 3. Optical flow calculation. (a) frame *i*, (b) frame *i+1*, (c) frame *i+2* and (d) frame *i+3*. Optical flow is calculated between (a) and (d). Results are represented in (e). Green vectors represent soft movements, yellow vectors for intermediate movements and red for strong movements.

the same patient may react differently during the same session), which made it necessary to analyze the motion in a global way, ruling out other classical solutions. To test the viability of the method, it was applied to different video sequences of real patients. Optical flow results were graphically represented to make easier their analysis.

For this experiment, we used full HD video sequences recorded with real patients during their audiometric evaluations. The global frame has a resolution of 1080x1920 pixels, allowing us to have enough resolution when we restrict the interest area to the eye region. Each audiometrical test takes between 5 and 8 minutes, but they can substantially increase their duration until 15 minutes or even more when the patients have severe cognitive impairment or when due to medications they are in a less reactive state.

In Figure 4 we can see several samples very representative of the optical flow results. These images show how the optical flow is able to correctly identify the movements produced between images. To see it more clearly, vectors were thesholded according to their strength representing in red the strongest movements, yellow for the intermediate movements and green for the softest. In Figures 4(a), 4(c) and 4(d) it is possible to appreciate a change in the gaze direction (specifically, the gaze is moved to the right). Particularly, in Figures 4(a) and 4(c) irises are located in the center of the eyeball for the first frame of reference, while for the second frame of reference they were moved noticeably at the right. According to this, red vectors (which represent the significant movements) are pointing to the right, correctly indicating the movement of the irises. In Fig. 4(d) the movement is the same but more subtle that in the two previous samples, nonetheless the optical flow is capable of detect it anyway. In the case of Figure 4(c) we can see the starting of an eye opening. In this case, the right eye opening is larger, so that more red vectors are located in this eye pointing upward. The same way, when there are no significant movements within the eye region, optical flow does

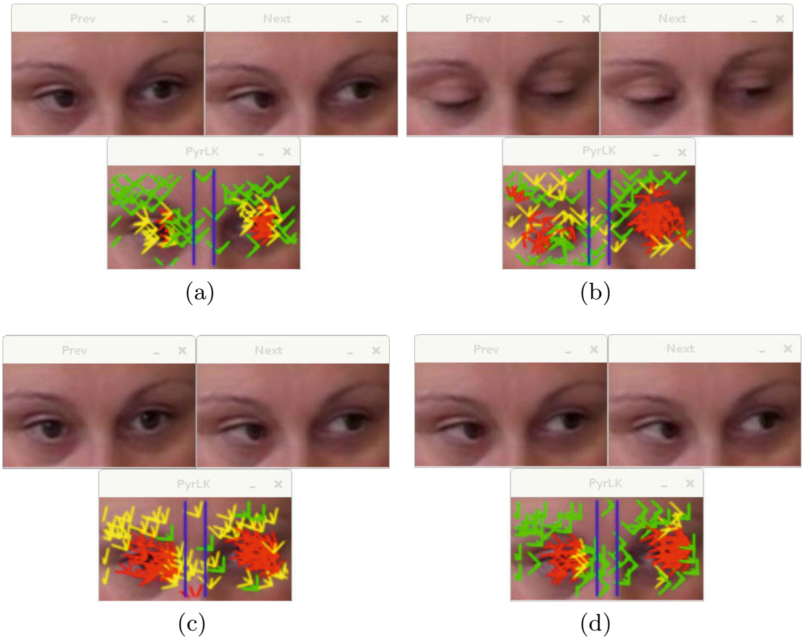

Fig. 4. Sample result images of movements and their corresponding optical flow vectors. Red vectors represent strongest movements, yellow for the intermediate movements and green for the softest. (a), (c) and (d) Change in the gaze direction to the right. (b) Eye opening.

not offer as a result any motion vector, or if it does it, they will be vectors of soft movements (those represented in green) not really significant movements (those represented in red).

All these four images are samples of the correct identification of the eye region movement by the optical flow proving its validity for this task. These results were also shown to the audiologists, whom corroborate them visually. This allow us to establish the optical flow as an initial detector of the movements that take places within the eye region. To give meaning to these vector we need to characterize them somehow and thus classify them as different movements that will be finally interpreted (or not) as a response to the auditory stimuli.

5 Conclusions and Future Research

In this work a preliminary solution for the detection of the movements within the eye region has been presented. The final aim of these researches is to support the audiologists when they are evaluating the hearing capacity of patients with cognitive impairments or other communication disorders. It is important to remember that each patient may present different gestures as a reaction, so movements must be analyzed in a global way. This preliminary study confirms the viability of the use of the optical flow for detecting and properly representing those movements. The results depicted in Section 4, along with many others that are not shown here but were also analyzed in detail, consistently represent the movements that the audiologist himself can appreciate.

As previously mentioned, future works will attempt to group these vectors and provide them with high-level information to identify them as expression changes and interpret them as a response to the auditory stimuli. To that end, we also need to get more video sequences when evaluating this particular group of patients. Because of their impairments many of these patients are entered in special centers and they do not come voluntarily to the consultation of the experts with whom we conduct this research. They are the experts who must move to visit these patients, but this requires special permissions. Recently we have obtained the necessary permissions to visit one of these centers and we have already evaluated all the patients entered there. These video sequences are going to be analyzed in next studies, which will allow us to corroborate our results and to advance the research.

The final contribution of this work could be very interesting for the audiologist community since it is a novel method for the global interpretation of the gestural reactions of this specific group of patients, which react in a way completely different to the standard patients. The evaluation of this particular type of patients has not been studied for its automatic resolution until now.

Acknowledgments. This paper has been funded by the Ministry of Science and Innovation of Spain (project TIN2011-25476) and by the Regional Ministry of Economy and Industry of the Xunta de Galicia (project 10TIC009CT).

References

1. Davis, A.: The prevalence of hearing impairment and reported hearing disability among adults in great britain. International Journal of Epidemiology 18, 911–917 (1989)
2. Espmark, A., Scherman, M.: Hearing confirms existence and identity-experiences from persons with presbyacusis. International Journal of Audiology 42, 106–115 (2003)
3. Fernández, A., Ortega, M., Cancela, B., Penedo, M., Vazquez, C., Gigirey, L.: Automatic processing of audiometry sequences for objective screening of hearing loss. Expert Systems with Applications 39(16), 12683–12696 (2012)
4. Viola, P., Jones, M.: Robust real-time object detection. International Journal of Computer Vision (2001)
5. Lucas, B.D., Kanade, T.: Iterative image registration technique with an application to stereo vision, vol. 2, pp. 674–679 (1981)