Gaze Alignment Techniques for Multipoint Mobile Telemedicine for Ophthalmological Consultations

Ramkumar Narayanan[1(✉)], Uma Gopalakrishnan[1], and Ekanath Rangan[2]

[1] Center for Wireless Networks and Applications (WNA), Amrita University, Amritapuri, Kollam, India
ramkumar@am.amrita.edu

[2] School of Medicine, Amrita University, Kochi, India

Abstract. Telemedical consultation systems are emerging as a viable medium for patient-doctor interaction in a number of medical specialties. Such systems are already prevalent in fields like cardio diagnosis and it is still very nascent in the field of ophthalmology. But with the emergence of affordable and high quality remote-control cameras, a host of new possibilities have opened up. In this paper, we have developed innovative gaze alignment techniques for ensuring Mutual Gaze, Gaze Awareness and Gaze following. The system is shown to work effectively even for interactions that are as complex as involving multiparty consultations involving remotely located patients through the use of a mobile telemedicine network and general physician/ physician-assistant and specialist ophthalmologist.

Keywords: Ophthalmology · Gaze · Mobile telemedicine

1 Introduction

There are several existing technologies that enable telemedicine consultation in the fields of cardiology [1, 2], neurology etc. However, in the field of ophthalmology, gaze is an important factor that directly impacts the effectiveness of an ophthalmological diagnosis. It is observed that out of the 45 million blind worldwide, 80% is curable. The vast rural populous, particularly in the developing world cannot afford visits to super specialty tertiary hospitals. Mobile Telemedicine Units on vehicles often drive to remote villages and setup camp to provide free checkup.

We have embarked on setting up of a multipoint mobile telemedicine network for ophthalmological consultation, consisting of an (1) ophthalmological ***specialist*** in a Tertiary Medical Center (TMC) located in an urban area, (2) a ***general practitioner*** in a Primary Health Care (PHC) center located within the county through which the care for all patients in the neighborhood is usually coordinated, and (3) a Mobile Telemedicine Unit (MTU) that reaches out to remote villages belonging to that county. In such a distributed multipoint telemedicine consultation scenario, targeted to address ophthalmological conditions, there is a need for the patients' gaze to be directed towards either the specialist or the general practitioner periodically during

P. Perego et al. (Eds.): MobiHealth 2017, LNICST 247, pp. 182–189, 2018.
https://doi.org/10.1007/978-3-319-98551-0_21

the course of the consultation, In fact, there are three major levels of gaze alignment [3, 4] that can be identified viz,.

(1) ***Mutual Gaze:*** simply refers to eye-contact between interacting patients and doctors.
(2) ***Gaze Awareness:*** which in this context means knowing where others are looking. The ophthalmological specialist often needs to perceive the gaze direction of the patient while performing the diagnosis.
(3) ***Gaze Following:*** which reflects an "expectation-based type of orienting in which an individual's attention is cued by another's head turn or eye turn". The ophthalmologist often points his/her fingers at a visual chart asking the patient to look at the object.

In this paper, we develop techniques to dynamically implement the above three levels of gaze alignment so as to transform a mobile telemedicine consultation into an effective medium for ophthalmological diagnosis. Our technique uses a media rich setup that dynamically maps appropriate camera feeds to display units such that gaze directionalities are preserved.

The system is being tested with a tripartite test-bed consisting of a patient at one location, a specialist in another and a general physician in the third location. The system is being tested in our setup consisting of Amrita Institute of Medical Sciences (a 1500 bed super specialty hospital) located in the city of Cochin in Southern India and Amrita Center for Wireless Networks located in Amritapuri, a picturesque rural village on the shores of Arabian Sea 100 km away. Feedback from preliminary experiments is promising.

2 Related Work

There are several telemedicine consultation systems such as MDLIVE, Teladoc, American Well, Doctor on Demand etc. However, they are mostly web portals coupled with peer to peer video conferencing systems. Such systems work fine for direction insensitive consultation systems.

Work done by Blackwell et al. [5], conducted an extensive survey in for ophthalmological telemedicine diagnosis. They conducted their experiments in the remote villages of Australia. This study was conducted as early as 1997 which showed that effectiveness of remote diagnosis did not quite match the face to face interaction.

Work done by Academy of Pediatrics Section on Ophthalmology - Pediatrics, 2015 [6], describes a retinal digital imaging technique for remote detection of retinal impairment.

In the field of head and eye pose detection, Sheela et al. [7], describes an iris based video tracking solution for estimation of gaze directions.

Bai et al. [8], describes tele-ophthalmology system for rural eye care systems such as Aravind Tele ophthalmology Network. These have been proven to be effective.

Ramkumar et al. [9, 10] describes some fundamental techniques in the area of gesture and gaze in an eLearning scenario. This work serves as the motivation for the application described in this paper and includes gesture triggered, gaze switching with the help of rich media devices.

3 Gaze Alignment Architecture in Ophthalmological Telemedicine

Ophthalmology requires a specialist to be able to direct the patient's gaze towards specific targets, such as reading charts or gaze at other objects. For instance, detecting and measuring degree of squint requires a system to provide the doctor with a frontal perspective of the patient. In another example, detecting impairment in peripheral vision in glaucoma patients may require the ophthalmologist to precisely perceive the patients' gaze direction.

Our proposed telemedicine architecture consists of three geographically separated locations viz.,

- Mobile Telemedicine Unit (MTU) hosting the patient - ***PA***.
- Primary Healthcare Center (PHC) consisting of a general practitioner – ***GP***.
- Tertiary Medical Center (TMC) consisting of a specialist (ophthalmologist) – ***SP***.

Let us take a typical consultation interaction pattern i.e., ***SP talks to PA***. *SP* takes the role of the speaker and *PA* takes the role of the listener. In fact, we define three possible roles in any consultation. We define ***role*** to be a temporary state the users are in depending on their involvement in the consultation interaction. We can observe three different roles in any consultation interaction namely,

(1) ***Speaker:*** is one of *SP*, *GP*, *PA* who is doing the talking in the consultation at that moment.
(2) ***Listener***: is the one to whom the speaker is mainly talking to.
(3) ***Observer:*** others who are passively witnessing the interaction.

As the consultation interaction proceeds, *PA*, *GP* and *SP* can take on different roles. An interaction is represented by $speaker \rightarrow listener$.

For this particular interaction in which *SP* talks to *PA*, represented by $SP \rightarrow PA$, let us derive the gaze directionalities. We notice the following in TMC (speaker's location), MTU (listener's location) and PHC (observer's location).

- At TMC, since *SP* is talking to *PA*, *SP* will gaze at the display that shows *PA* (denoted by D_{PA}) and this is represented by a vector called the ***entity gaze vector***, $\overrightarrow{SP \rightarrow D_{PA}}$ (shown using green arrow in Fig. 1a).

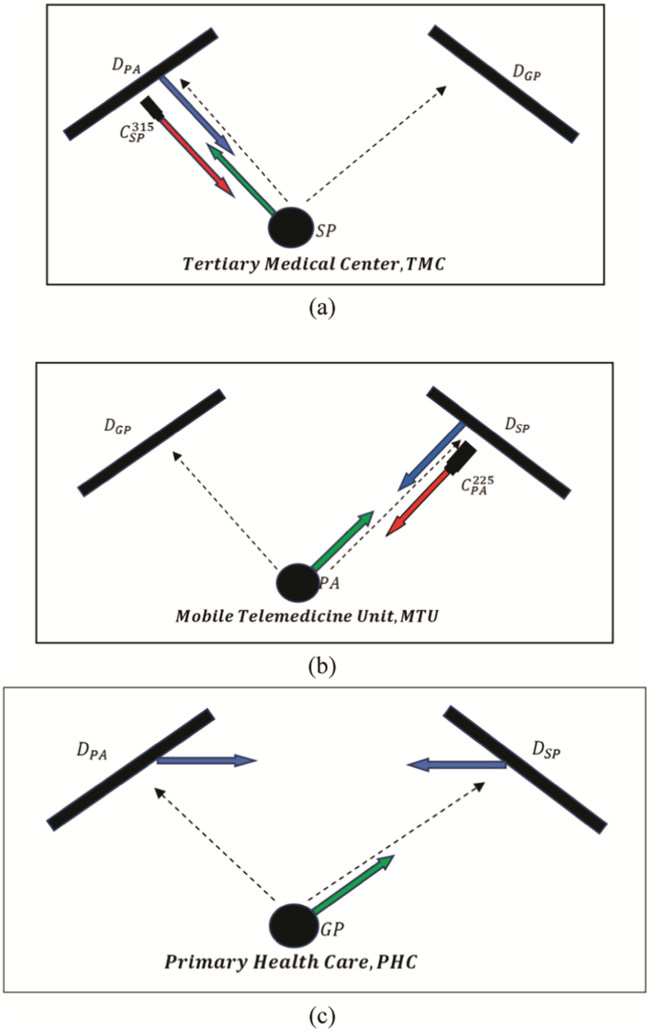

Fig. 1. Ophthalmological consultation consisting of *SP* talking to *PA*. Gaze alignment in (a) TMC, (b) MTU and (c) PHC. (Color figure online)

- At MTU, *PA* will look at the display showing *SP* (denoted by D_{SP}) and his/her entity gaze vector is given by $\overrightarrow{PA \rightarrow D_{SP}}$ (shown using green arrow in Fig. 1b).
- At TMC, to ensure eye contact, we require $PA's$ gaze direction as displayed on D_{PA} to be towards *SP*. This is given by the ***display gaze vector***, $\overrightarrow{D_{PA} \rightarrow SP}$ (shown by a blue arrow in Fig. 1a).
- Angle between $\overrightarrow{D_{PA} \rightarrow SP}$ and $\overrightarrow{SP \rightarrow D_{PA}}$ in TMC should be 180° to enable mutual gaze. Now go into MTU, take the reverse of this angle (−180°) from the entity gaze

vector, $\overrightarrow{PA \rightarrow D_{SP}}$. This yields the ***camera gaze vector***, $(\overrightarrow{C_{PA}^{225} \rightarrow PA})$ which is directed from the camera, C_{PA}^{225} towards *PA* in Fig. 1b, denoted by a red arrow.

- In order to arrive at the camera gaze vector in TMC, we follow a similar approach with a view to present $SP's$ gaze displayed on D_{SP} in MTU, so as to appear as though directed towards *PA*. This will yield the camera gaze vector as shown in Fig. 1c, denoted by the red arrow.
- The gaze directionalities of the observer, *GP* in PHC presents itself with a bit more complexity. The gaze direction of *GP* is towards the display showing the speaker, D_{SP}. The gazes of D_{SP} and D_{PA} should appear directed towards each other as represented by blue arrows in PHC. How does this translate into camera gaze vectors at PHC? This is derived from the generalized gaze mapping algorithm that we present in the next section.

4 Generalized Gaze Mapping Algorithm

When we generalize the interaction algorithm, there can exist six consultation interaction patterns, viz., $SP \rightarrow PA$, $PA \rightarrow SP$, $SP \rightarrow GP$, $GP \rightarrow SP$, $PA \rightarrow GP$ and $GP \rightarrow PA$, all of which are handled by the generalized gaze mapping algorithm presented below.

Configuration Steps:

Step 1: Map *entities*, $E \in \{SP, GP, PA\}$ to roles, $R \in \{speaker, observer, listener\}$.
Step 2: In each of the three locations, TMC, PHC, MTU, the displays are mapped to distinct remote entities and are called D_{SP}, D_{GP} and D_{PA}. If and when we refer to displays as $D_{speaker}$, $D_{listener}$ and $D_{observer}$ these roles may be substituted by the respective entites.

Entity Gaze Vectors (all drawn in green):

Step 3: The entity gaze vectors of the listener and observer in their respective locations, is directed towards display showing the speaker.
Step 4: In the speaker's location, his/her entity gaze vector is directed towards the display mapped to the listener.

Display Gaze Vectors (all drawn in blue):

Step 5: In the listener's location, a display gaze vector is drawn from the $D_{Speaker}$ towards the *listener*.
Step 6: In the listener's location, a display gaze vector is drawn from the $D_{Observer}$ towards the $D_{Speaker}$.
Step 7: In the observer's location, a display gaze vector is drawn from the $D_{Speaker}$ towards the $D_{Listener}$ and vice versa.
Step 8: In the speaker's location, two display gaze vectors are drawn from the $D_{Listener}$ and $D_{Observer}$ towards the speaker.

Relative Position Vectors, (all drawn in dotted black):

Step 9: Relative position vectors are drawn from an entity to each of the displays in the entity's location.

Computation of Camera Gaze Vectors, (all drawn in red):

Step 10: The camera gaze vector directed towards entity Ei, whose video is presented on a display to entity Ej, where $Ei \neq Ej \in \{SP, GP, PA\}$ is to be computed as follows,

Substep 10.1: calculate θ = angle between

$$relative\ postion\ vector, \overrightarrow{Ej \rightarrow D_{Ei}}$$

and

$$display\ gaze\ vector,\ \overrightarrow{Ei\ in\ Ej's\ room}$$

Substep 10.2: $camera\ gaze\ vector = Ei's\ entity\ gaze\ vector - \theta$

Applying the above algorithm for each of the six interactions, we arrive at the table of angles of camera gaze vectors. Table 1 enumerates them. The notation uses C^{φ} to indicate a camera positioned at angle φ from the x axis.

Table 1. Angles of camera gaze vectors obtained by applying the generalized gaze vector algorithm for each of the interaction consultations. Camera C^{φ} indicates a camera oriented at an angle, φ to capture E_i's video in E_i's location and this video is routed to appropriate display at E_j's location so that E_j gets to view a gaze aligned perspective of E_j. Camera locations are highlighted in red whereas display locations are highlighted in blue.

Interactions (Speaker → Listener)	$Ei = PA$ (in MTU) $Ej = SP$ (in TMC)	$Ei = PA$ (in MTU) $Ej = GP$ (in PHC)	$Ei = GP$ (in PHC) $Ej = SP$ (in TMC)	$Ei = GP$ (in PHC) $Ej = PA$ (in MTU)	$Ei = SP$ (in TMC) $Ej = PA$ (in MTU)	$Ei = SP$ (in TMC) $Ej = GP$ (in PHC)
$SP \rightarrow PA$	C^{225}	C^{270}	C^{225}	C^{180}	C^{315}	C^{0}
$PA \rightarrow SP$	C^{225}	C^{270}	C^{270}	C^{315}	C^{315}	C^{0}
$GP \rightarrow SP$	C^{270}	C^{270}	C^{225}	C^{180}	C^{270}	C^{225}
$SP \rightarrow GP$	C^{225}	C^{180}	C^{225}	C^{180}	C^{270}	C^{225}
$GP \rightarrow PA$	C^{270}	C^{315}	C^{0}	C^{315}	C^{270}	C^{225}
$PA \rightarrow GP$	C^{270}	C^{315}	C^{0}	C^{315}	C^{315}	C^{0}

As the consultation moves from one interaction to another the appropriate video switching subsystem chooses the camera corresponding to the appropriate row (determined by the interaction) and column (determined by the camera and display locations - shown in red and blue respectively) in Table 1. A fully equipped telemedicine consultation room with all of the cameras and displays installed is shown in Fig. 2.

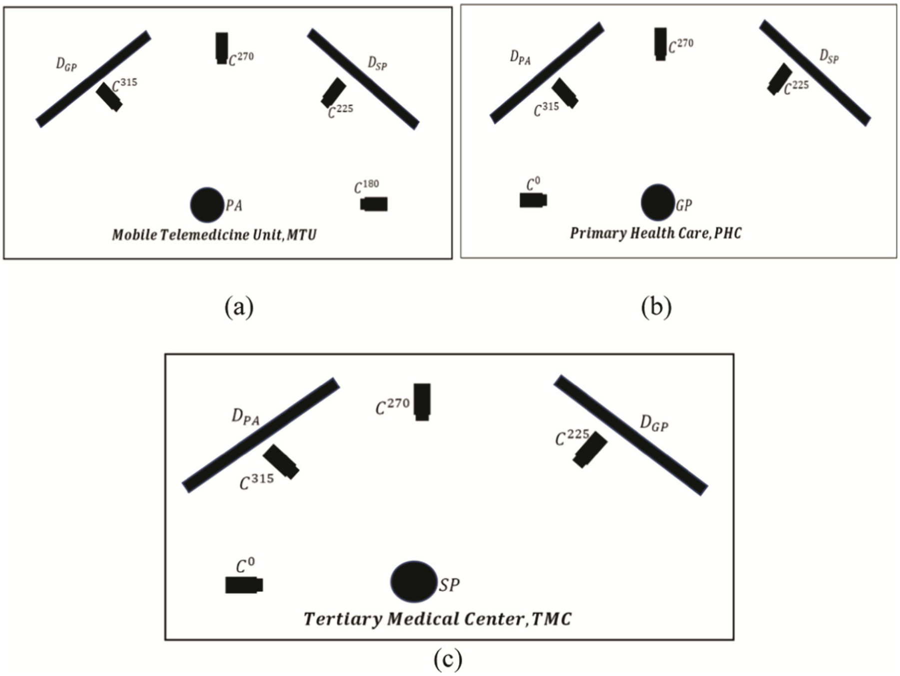

Fig. 2. A fully equipped telemedicine consultation room with gaze alignment for all consultation interaction patterns (a) MTU, (b) PHC and (c) TMC

5 Prototype Implementation

We have implemented a prototype of our system and conducted sample sessions with participants who took on the various roles (see Fig. 3). For instance, the ophthalmologist was asked to detect squint eyes and conduct a basic test for strength of peripheral vision (suspecting glaucoma) by directing the gaze of the patient at certain targets indicated by

Fig. 3. A prototype system in action for a gaze aligned interactive consultation.

gestures. The ophthalmologist was able to carry out the consultation with much greater ease and naturalness.

6 Conclusion

We have implemented a gaze alignment system for multipoint mobile telemedicine consultation for ophthalmological diagnosis. Since the field of ophthalmology has a direct connectedness to gaze and directionalities in general, this technology is a niche fit. However, it can be extended to other areas of medical diagnosis which are interaction intensive and requires gaze alignment. A prototype implementation shows promising initial results.

References

1. Pathinarupothi, R.K., et al.: RASPRO: rapid summarization for effective prognosis in wireless remote health monitoring. In: Wireless Health (2016)
2. Pathinarupothi, R.K., Rangan, E.: Discovering vital trends for personalized healthcare delivery. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. ACM (2016)
3. Monk, A.F., Gale, C.: A look is worth a thousand words: full gaze awareness in video-mediated conversation. Discourse Process. **33**(3), 257–278 (2002)
4. Flom, R., Lee, K., Muir, D. (eds.): Gaze-Following: Its Development and Significance. Lawrence Erlbaum Associates Publishers, Mahwah (2007)
5. Blackwell, N.A.M., Kelly, G.J., Lenton, L.M.: Telemedicine ophthalmology consultation in remote Queensland. Med. J. Aust. **167**(11), 583–586 (1997)
6. Fierson, W.M., Capone, A., American Academy of Pediatrics Section on Ophthalmology.: Telemedicine for evaluation of retinopathy of prematurity. Pediatrics **135**(1), e238–e254 (2015)
7. Sheela, S.V., Abhinand, P.: Iris detection for gaze tracking using video frames. In: 2015 IEEE International Advance Computing Conference (IACC). IEEE (2015)
8. Bai, V.Thulasi, et al.: Teleophthalmology-based rural eye care in India. Telemedicine e-Health **13**(3), 313–321 (2007)
9. Ramkumar, N., et al.: Gesture triggered, dynamic gaze alignment architecture for intelligent eLearning systems. J. Intell. Fuzzy Syst. **32**(4), 2963–2969 (2017)
10. Narayanankutty, R., et al.: Automatic multi-perspective switching for gaze alignment in e-Learning systems. IJAER **10**(16), 37303–37310 (2015)