

Geometric Estimation of FoV for Vision Modeling in DHM

Vinayak and Dibakar Sen
Indian Institute of Science

Copyright © 2009 SAE International

ABSTRACT

This paper presents a geometric method for estimating the filed-of-view (FoV) for monocular and binocular vision using a 3D scanned human face. The traditional vision-cone approach offers simplicity but may lead to spurious inferences during ergonomic evaluations involving peripheral vision. FoV is asymmetric due to unrestricted zygomatic vision and restrictions on nasal side. Thus its accurate modeling is important for effective simulation of visual performance. Given a tessellated face model and the location of the pupil, the present work computes the FoV edges on the face with respect to the pupil. The algorithm is divided in two steps, visible triangle computation and silhouette edge detection for FoV estimation. A qualitative comparison of samples with the standard perimetry data and right circular cones used in DHM shows that the FoV is very similar to the physiological visual-field maps and the cones used in DHM are conservative. The proposed approach has been extended to gaze dependant field-of-view computation which is difficult to obtain experimentally due to involuntary eye movements in subjects. The gaze-dependant FoV is found to be significantly different from normal monocular and binocular FoV. Thus the present computational FoV paradigm has the potential of active vision integration in performance evaluations using DHM's.

INTRODUCTION

VISUAL PERCEPTION AND PERIMETRY - Vision is one of the five senses through which human beings perceive the geometric attributes (viz. shape, size, texture) of objects in their environment. Visual

perception in literature has been studied from anatomical, physiological [1] and psychological [2] [3] points of view. As an anatomical phenomenon, human vision deals with the incidence of a ray of light from a source on to the human retina which consists of the receptor cells (rod cells and cone cells). The portion of space in which objects are visible at the same moment during steady fixation of gaze in one direction [4] is referred to as the field-of-view (FoV). This field-of-view, when processed by the visual system, computes the spatial array of visual sensations available to observation and is referred to as the visual field [5]. The density of the receptor cells is more at the retinal centre (fovea) and varies across the span of the retina [5] (figure 1). This variation causes difference in the visual perception of objects depending upon their location in the FoV. Thus, human visual field is divided into two distinct but contiguous regions namely central (foveal) and peripheral. The central region has strongest visual acuity. Though the peripheral region has reduced visual acuity, varying with the distance from the central region, it has significant motion sensitivity.

The visual field is determined experimentally through the standardized procedure of perimetry [4]. Traditional perimetry involves a subject to gaze at a specified direction and respond to visual stimuli generated manually by the ophthalmologist. Recent literature in perimetry has proposed automated methods, like Goldmann visual field exam, Humphrey field exam and Tangent screen exam [4, 6], to obtain the visual field of the subjects. Such methods generally involve dome-shaped screens wherein the subject is required to respond to the computer generated stimuli on the screen.

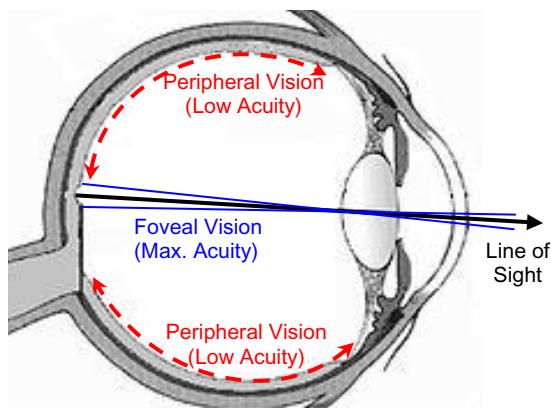
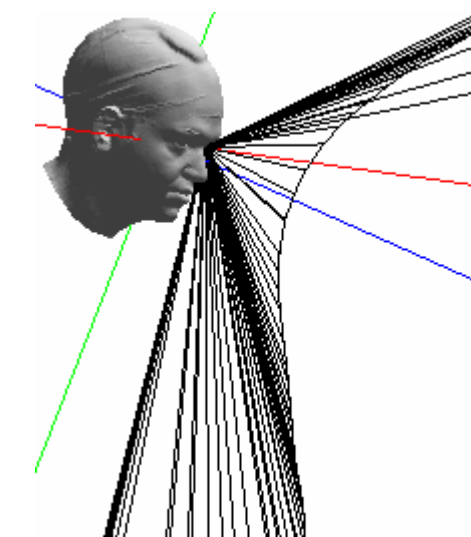


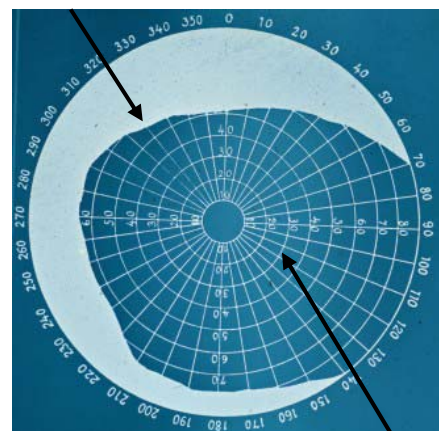
Figure 1: Human Visual System

The stimuli can be static or kinetic based on the control of stimuli by the computer. These methods apparently provide better accessibility, repeatability and accuracy. However, since the procedures involve interactive response by the naïve subjects and possibility of involuntary gaze variations, the effect of these human errors on the repeatability and accuracy of the results needs investigation. These procedures are inherently time-consuming and demand repetition for reliable readings. This has obvious implication on the resources. In addition, visual field tests focus more on the determination of pathologies within the FoV of a given patient rather than statistical variation of FoV across different subjects which is more relevant to ergonomics applications and DHM. Possibly, for this reason the accuracy of the extent of FoV has not been the subject matter of much discussion in medical literature. In ergonomics literature, the effect of FoV on user performance is well-addressed for assessment of pilot head movements with respect to FoV [7], ergonomics evaluation of virtual cockpit environments [8], and evaluation of judgment based measures for designing synthetic vision systems [9]. Further, the combination of ergonomics and FoV analysis has also been mentioned Rix and Stork [10]. Despite this, accurate determination of FoV has not received much attention in these areas. Although it is mentioned Kroemer *et. al.* [11] that the visual field and hence the FoV is primarily affected by the anthropometry of human face (figure 2(a)), specific studies on this could not be found by the present authors.

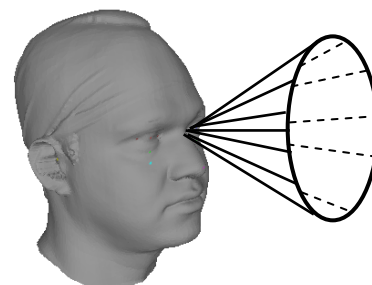
VISION AND HUMAN PERFORMANCE - Natural human performance during object manipulation involves four distinct activities namely search (intelligent sampling of visual field for presence of the object of interest), locate (estimation of position of the object in the environment or relative disposition of the object with respect to other objects in the scene), reach (distance minimization, obstacle avoidance, target fixation, tool/hand tracking to target) and execute (inspection). For example, let us consider the task of *opening a screw* in light of the four above mentioned stages. Given this task, the worker *samples* the visual field and *compares* the objects in the visual field to the screw.



Actual Visual Field (a)



(b) 30° Cone



(c)

Figure 2: Actual and Modeled FoV (a) asymmetric human visual field, (b) perimetry vs. right cone – polar map¹ (c) right-circular vision cone (3D)

The context, experience and knowledge provide the worker with an estimate of probability with which the object of interest can be found at a given location [12] [13]. So, the worker *searches* for the object of interest in selected portions of the environment where the

estimated probability is high rather than scanning the whole space. This makes the sampling effective, especially with the visual system, which provides one with only a selected subset of objects in the scene whose size and details in the visual field are naturally graded; nearer objects are fewer with more details and vice versa. This helps one to maintain the attention only within a limited region. Thus the FoV plays an important role in object search. Once any sampled object is identified as the screw, the worker estimates the position of the screw. This estimation is about where the screw is relative to the other objects such that a tentative path to reach the screw can be planned based on his visual understanding of the object clutter. With the target fixated, the worker searches and procures a suitable tool – the screw-driver. This is done by moving the head to keep the target in central vision. The decision on the kind of screw-driver to be used is generally based on the visual cues taken from the head of the screw. The *reach* activity involves taking the tip of the screw-driver to the head of the screw minimizing the distance between the tool and the object of interest. While reaching the object, the worker uses the planned path to the destination which can get altered to avoid collision with obstacles. This activity involves the worker to track the motion of the tool while his *gaze is fixated* on the screw. By virtue of peripheral vision, the movement of the tool to the screw can be easily detected till the tool comes in central vision. Finally, for opening the screw a visual estimation of the position and orientation of the slot helps the worker to orient the tip of the screw-driver. While rotating the screw-driver, the motion of the screw outward to its location of fixture acts as a feedback to the worker. The rotation of the tool for opening the screw is also visually inspected to verify the desired outcome of the task i.e. the disassembling of the screw. It is seen here how vision and the FoV play a central role in a man-machine system while manipulating objects of interest.

The behavioral modeling of DHM has been the focus of research [14, 15, 16, 17] for providing naturalness to movements during object manipulation in virtual environments. To address the emulation of natural human behavior, the existing literature has focused on motion correction [17] and motion planning algorithms for navigation [15]. Literature in cognitive sciences and DHM [3, 18] establishes the importance of vision understanding to create realistic models of the “search” and “locate” activities. To this end, the existing DHM research has also concentrated on incorporating vision based techniques like visibility-based probabilistic roadmaps [19] for motion planning and vision-performance for posture prediction [20].

On the other hand, in VR environment the visual frustum used for view clipping in computer graphics is itself used as an estimate of users FoV for stereoscopic rendering of the virtual scene on to the HMDs. The view so generated has been referred to as Geometric Field of View (GFoV) [21, 22, 23]. Although the relation and

relevance of GFoV and human FoV is not clear, it can be observed that even in VR based studies involving real human interacting with virtual environment, the role of accurate modeling of the FoV, for appropriate evaluation of human performance through simulation studies, has been emphasized. In view of the role of FoV in evaluation of tasks involving visual performance, effective and reliable determination of FoV for normal (non-pathological) subjects is necessary.

VISION CONE MODELING IN DHM - Traditionally, DHM has used 30° and 45° right circular cones at the pupil as FoV's [14] to represent central and peripheral vision respectively. Although this model offers simplistic solution for the object-visibility problem, actual ergonomics analysis of the situation is done by experts through the study of graphical rendering. More realistic visibility rendering would surely aid the experts in making better evaluations. Certain ergonomics evaluations such as hand-eye coordination [24], placement of blinking light indicators in an aircraft cockpit environment, involve motion detection which is better assessed by peripheral vision. The standard perimetry maps (figure 2(b)¹) suggest that vision on the zygomatic side of the eye is largely unrestricted. Thus, the 45° limit to the peripheral FoV might lead to spurious inferences by reducing the extent of peripheral vision significantly (figure 2(c)). The effect of face anthropometry being a major factor in FoV asymmetry [11], it is important to make use of face geometry to estimate FoV reliably for ergonomic evaluations involving vision as a critical component. This paper presents a computational geometric approach to estimate the FoV from the 3D scanned tessellated models of human faces.

DATA ACQUISITION AND LAND MARKING

Heads of 50 subjects wearing bald-caps with eyes open were scanned using a Cyberware® Head and Face 3D color Scanner and stored in .ply format containing information about vertices, indexed triangles and color of each vertex on the face [25]. Originally each tessellated face contained 100,000 to 150,000 triangles. The scanned model was smoothened, rectified and data-decimated using the mesh-processing software MeshLab™². The resulting face model contained about 40,000 triangles.

Different positions and orientations of heads with respect to the perimeter-screens would give maps which are apparently different and difficult to compare. Thus in perimetry, a standard frame of reference is necessary for positioning and orienting the head appropriately. As

¹ University of West Indies, Mona:
<http://www.mona.uwi.edu/fpas/courses/physiology/neurophysiology/Perimetry.htm>

² MeshLab Visual Computing Lab-ISTI-CNR
<http://meshlab.sourceforge.net/>

per the standard, the head is required to be horizontal with the face pointing to the perimeter-screen.

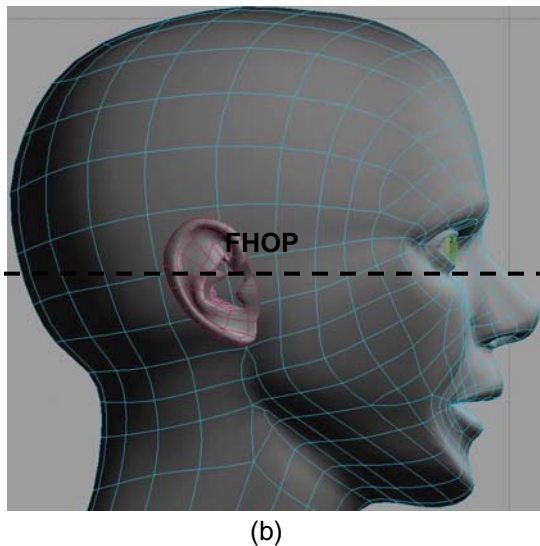
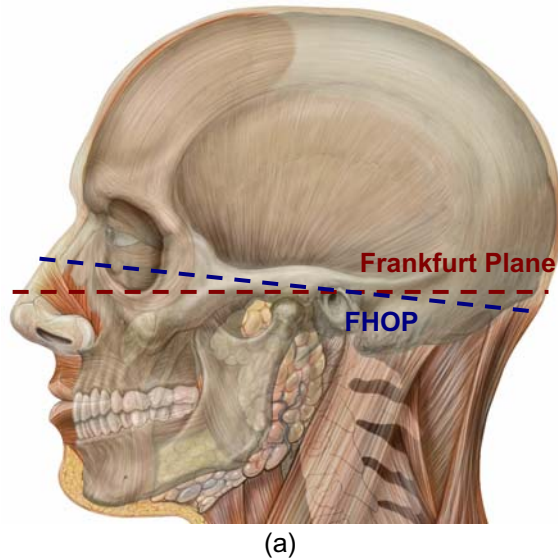


Figure 3: Head orientation planes (a) skeletal head orientation (Frankfurt) plane², (b) proposed facial head orientation plane (FHOP)³

Further, the head is positioned to facilitate the subject with gaze fixation at the desired location on the screen. Present ergonomic standards consider the frontal, sagittal and Frankfurt planes as the three mutually perpendicular planes defining the coordinate frame of the face [11]. The Frankfurt plane, used for aligning the head with the horizontal, is defined as the plane determined by the highest point on the upper margin of the opening of each external auditory canal and the low point on the lower margin of the left orbit (figure 3(a)³). It may be noted that the Frankfurt plane is defined on the skull. On a live subject the lower margin of the left orbit is identified by palpation by experts. This is not possible for scanned face models as the lower orbital is not

geometrically prominent. Thus the Frankfurt Plane is not accurately identifiable in a purely geometric approach.

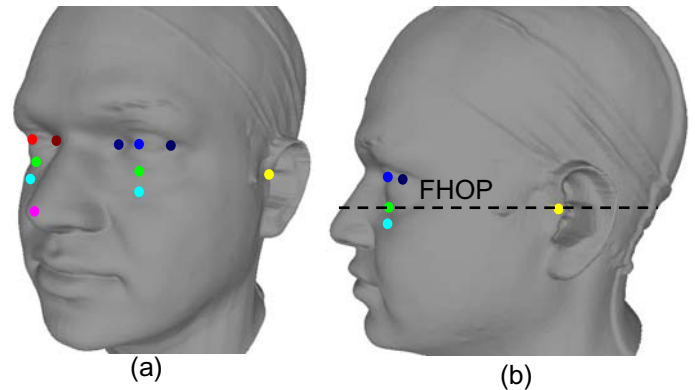


Figure 4: Head model preparation, (a) landmarks on the scanned head model, (b) facial head orientation plane

It is observed that in the vicinity of the lower orbital, lower eye-lids are relatively immobile for different states of the ocular fissure; even during blinking mostly the upper eyelid moves. Hence, using the base of the two lower eye-lids (figure 3(b)⁴) and the top edges of the two ear canals, an alternate plane, referred to here as the Facial Head Orientation Plane (FHOP), has been defined and used in the present work to standardize the position and orientation of the face scan with respect to the global frame of reference. The face model was first colored with a single color (grey) and thirteen selected facial landmarks (figure 4(a)) were interactively marked with color codes.

Work on automated face anthropometry by Deo [26] can be used to identify the above mentioned landmarks on the face but it requires robust algorithms for assuring the reliability of the face landmarks. For the present work, face-landmarks have been identified and marked manually. The FoV plane is fit using the ear and eye-lid landmarks and the nose-tip identifies the front direction of the face (the one parallel to the Z-axis). It can be seen that the FoV plane defined by the ears and eye-lids is an acceptable approximation for the Frankfurt plane (figure 3(b) and figure 4(b)). These five points taken together are used for placing and orienting the face model on the global frame of reference. The role of the eye landmarks is to identify the points with respect to which visibility is computed as will be discussed in the following sections.

COMPUTATION OF FOV

The portion of space in which objects are visible at the same moment during steady fixation of gaze in one direction [4] has been defined as the field-of-view (FoV) in perimetry literature. From geometric optics point of view FoV can be conceived as the set of all rays that

³ P. J. Lynch, medical illustrator; C. C. Jaffe, MD, cardiologist: <http://creativecommons.org/licenses/by/2.5/>

⁴http://blog.mathieudesuche.net/public/Face_modeling/M_y_face_profile_mesh2.jpg