[5] I. Sabuncuoglu and D. L. Hommertzheim, "Dynamic dispatching algorithm for scheduling machines and automated guided vehicles in a flexible manufacturing system," *Int. J. Product. Res.*, vol. 30, pp. 1059–1079, 1992.

[6] D. Yim and R. J. Linn, "Push and pull rules for dispatching automated guided vehicles in a flexible manufacturing system," *Int. J.Product. Res.*, vol. 31, pp. 43–57, 1993.

[7] Z. Xiaobo, K. Ohno, and K. Nakade, "Modeling for flexible manufacturing systems and structural properties of optimal work routing policy," *J. Intell. Manufact.*, vol. 8, pp. 497–503, 1997.

[8] D. D. Yao and J. A. Buzacott, "Models of flexible manufacturing systems with limited local buffers," *Int. J. Product. Res.*, vol. 24, pp. 107–118, 1986.

[9] L. Zhuang and K. S. Hindi, "Approximate decomposition for closed queueing network models of FMS's with a block-and-wait and state-dependent routing mechanism," *Eur. J. Oper. Res.*, vol. 67, pp. 373–386, 1993.

[10] M. L. Puterman, *Markov Decision Process*. New York: Wiley, 1994.

[11] Z. Xiaobo, K. Ohno, and K. Nakade, "An optimal part routing policy in a flexible manufacturing system with an FMS blocking mechanism and a BDSM part routing," *IIE Trans.*, to be published.

Rotation and Direction Judgment from Visual Images Head-Slaved in Two and Three Degrees-of-Freedom

Bernard D. Adelstein and Stephen R. Ellis

Abstract—The contribution to spatial awareness of adding a roll degree-of-freedom (DOF) to telepresence camera platform yaw and pitch was examined in an experiment where subjects judged direction and rotation of stationary target markers in a remote scene. Subjects viewed the scene via head-slaved camera images in a head-mounted display. Elimination of the roll DOF affected rotation judgment, but only at extreme yaw and pitch combinations, and did not affect azimuth and elevation judgment. Systematic azimuth overshoot occurred regardless of roll condition. Observed rotation misjudgments are explained by kinematic models for eye-head direction of gaze.

Index Terms—Fick gimbal, Listing's law, spatial situation awareness, telepresence, virtual environment.

I. INTRODUCTION

In telepresence or virtual environment (VE) applications, the display of correctly positioned and oriented images potentially requires that the kinematics of the viewing system be able to conform to the user's motion capability. In a VE, the observer's gaze direction, range, and offset derived from sensor readings are geometrically transformed by the computer to render spatially stable images in the presence of head motion. In a teleoperation scenario, image stability is maintained by computer-controlled actuators that drive cameras on a remote platform to match the observer's head movements.

Full tracking of the observer's viewpoint may therefore necessitate monitoring all kinematic degrees-of-freedom (DOF's) associated with head and body movement. However, the higher the number of DOF's involved in head tracking and rendering the displayed scene, the greater

the hardware and software complexity will be. Increased complexity, in turn, entails possible performance degradation and higher equipment cost.

Allowing for more DOF's adds to the internal computation demanded of motion sensing and image rendering equipment. If motion measurement employs kinematic linkages coupled to the human, additional structure and sensors are required for tracking more DOF's. These hardware problems are further compounded in the design of head-slaved camera platforms, where more complicated linkages require not only sensors and structure, but actuators as well. A result of the added structure and actuators is degradation in the dynamic response and accuracy of camera platform control. Consequently, there is practical interest in designing systems with minimum numbers of DOF's.

This experimental study examines the kinematics of telerobotic camera platform "heads"—devices built specifically to replicate the orientation capability of the human head with its specific function of directing gaze. In humans, deliberate (other than artifactual due to head-neck rotation biomechanics) head translation arises mainly from motion of the torso. Similarly, in telerobotic systems, translation is generated by the platform's base which may be embodied by a mobile vehicle or an anthropomorphic torso. Thus, we investigate whether the kinematic complexity associated with the addition of the third orientation DOF to the typical (cf. [1, p. 199]) "pan and tilt" camera platform benefits operator spatial situation awareness as quantified by direction and rotation judgments. We focus on direction and rotation because these are the visual display components most significantly impacted by the projective geometry of changes in head orientation. The consideration of whether to include a third roll DOF has similar implications both for head motion sensing and for rendering VE simulations.

Below, we first review the kinematic architectures of relevant camera platforms. This summary motivates a remote viewing experiment on the judgment of the direction and rotation of static target objects in a simple environment. Experimental data for the effect of two and three DOF camera platform configurations on subject direction and rotation judgments are then reported. Discussion of these results in terms of kinematic motor behavior models for coordinated eye-head direction of gaze follows.

II. CAMERA "HEAD" KINEMATICS

Several camera platforms built specifically for head-slaved telepresence applications ([2]–[5]) have offered kinematics with more than the two DOF's of commercially available pan-tilt surveillance devices. The three DOF stereo camera platform built by Bolas and Fisher [5] used in the experiments described below has a kinematic configuration replicating that of one of the earliest head-slaved systems [2]. Both of these platforms have a motorized pan (yaw) base that carries a tilt (pitch) DOF to point cameras, with a geared differential arrangement between the pair of tilt motors to provide a third twist (roll) DOF. This sequence of a fixed vertical yaw axis at the base supporting a pitch axis moving in the horizontal plane corresponds to the Fick gimbal configuration shown in Fig. 1. The addition of the third twist axis orthogonal to the pitch axis of the two DOF Fick gimbal forms one convention in the oculomotor literature for describing three-dimensional (3-D) orientation [6, p. 181].

Our three DOF platform has a spherical center of rotation (i.e., the intersection point of all three rotation axes) located behind and 50 mm below the two camera's optical entrance pupils, a feature which Bolas and Fisher considered to approximate human eye and head biomechanics. The effect of this rotational center offset is that the platform
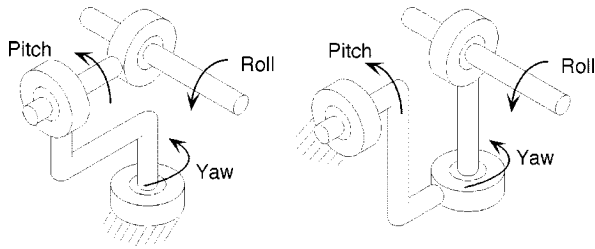
Fig. 1.   Fick gimbal (left) and Helmholtz gimbal (right).

pivots below at a "neck," causing translation at the cameras concomitant with head pitch and roll angle changes. The device's inter-camera separation and vergence are not motorized and must be preadjusted manually.

The inclusion of a third roll DOF was motivated by Bolas and Fisher's preliminary evaluation of a two DOF Fick gimbal camera platform design in a VE mock-up within which computer generated imagery of a telepresence situation could only be oriented in response to the user's head pitch and yaw motion [5]. In the absence of response to head roll, users complained of difficulty locating objects off to the side or "over the shoulder" and of an increased sensation of "queasiness" due to unsteadiness of their visible horizon.

An obvious explanation of these effects is the inability of a two DOF Fick gimbal (pan-tilt) platform mechanism to follow arbitrary 3-D head orientations. This inability resulted in images that did not conform to the users' internally perceived head orientation. Tweed and Vilis [7] and Straumann *et al.* [8] reported that, when the head and eye are free to move during either target-directed or random gaze, orientation changes for both the head and eye relative to fixed space follow two DOF "Listing's law" trajectories [6]. In a Listing system, the 3-D axis of rotation for any orientation change is constrained to the plane perpendicular to the fixed "primary position"—nominally, the "straight ahead" direction. Thus, for a given gaze direction, the two DOF Fick gimbal will have an identical pointing direction but a different twist (i.e., roll angle) than the corresponding Listing orientation. Glenn and Vilis [9] refined these observations to include larger gaze deflections and showed that orientations of the head in space were *Fick-like*, having twist intermediate to the Fick and Listing systems. Glenn and Vilis further pointed out that twist differences between the Fick and Listing systems become significant at large eccentricity, oblique gaze directions—i.e., large angular excursions that combine both yaw and pitch motions away from the primary position.

In this paper and elsewhere [10], we examine how head-slaved camera platform *orientation* kinematics affect an observer's ability to judge the direction and rotation of objects in a remotely viewed scene. Direction and rotation are the important factors here because, geometrically, these are the spatial displacements directly produced by arbitrary, quasi-static, three DOF orientation changes. Of interest are platform configurations which might be incapable of tracking fully all orientation DOF's in an observer's head movement—in this case, by specifically disabling the roll DOF in a Fick gimbal system. Since we are concerned here only with judgments of direction to the remote objects and rotation of these objects about their direction vector, distance cues that would arise from actively modulated camera vergence or from motion parallax were minimized by the design of the experiment.

III. METHODS

A. Apparatus

The experimental apparatus, shown in Fig. 2, is divided into a remote station containing the camera platform and a local station at which the

subject is seated. Both stations consist of identical 1.2 m tall, 0.61 m radius, partial (225° circumference) cylinder FoamCore taskboards. The camera platform is located such that its yaw axis coincides with the remote station's cylindrical axis and its cameras point at the cylinder's inner surface. The subjects were similarly positioned during the experiments, facing the inside of the local taskboard, with the taskboard's bottom edge at midthorax height.

The camera platform (Molly™ prototype, Fake Space Labs, Mountain View, CA) supports two parallel-mounted (75 mm separation between optical axes) miniature monochrome CCD cameras [5]. The platform's pan, tilt, and twist DOF's, were driven by computer controlled servo motors under orientation commands from an electromagnetic spatial sensor (Space Navigator, Ascension Technology Corp., Burlington, VT) affixed to the head-mounted display (HMD) (EyePhone, VPL Inc., formerly of Palo Alto, CA) worn by the subject. Our software on the supervisory personal computer (IBM-AT with 8 MHz 80286 CPU) and the camera platform's motor control cards (Model SMCC, Delta Tau Data Systems, Canoga Park, CA) read the spatial sensor's orientation and updated servo commands at a minimum 30 Hz rate. Since images were generated by the cameras and mechanically oriented by platform motion, image rendering on its own imposed no computational burden in this telepresence implementation. Delay in system response was predominantly attributable to the spatial sensor's ∼75 ms combined internal processing and serial line communication latency (slightly slower than the Ascension Bird reported in [11]) interposed between head orientation measurement and servo controller update. This delay was not considered an important factor in these experiments because of the quasi-static nature of the judgment task described below.

The HMD's two LCD screens received images transmitted from the corresponding left and right cameras. Each LCD had a resolution of 208 by 139 color triad "pixels." However, since the cameras were monochrome, images were rendered in black and white. The HMD was equipped with wide field of view (FOV) optics (LEEP Pop-Optix,Waltham, MA). Inverse lenses (also made by LEEP), which correct for the optical distortion in the HMD's optics [12], were mounted on the platform cameras. The HMD had a 75° horizontal by 60° vertical FOV across the center axis of each eye with a combined binocular horizontal FOV of 90° [12]. This horizontal and vertical FOV, however, is reduced at oblique viewing angles by the vignetting of the HMD lenses' cutout shape and underlying LCD mask.

While wearing the closed HMD, the subjects could view only the remote taskboard; they saw neither the local taskboard nor their own limbs. This paradigm for pointing at visual targets is termed "full open loop" (e.g., [13]). Open loop judgment is a normal requirement for visualization and exploration tasks where direct "hands-on" or remote manual interaction is absent. In the presence of time delay, open loop manual aiming is a necessity because the human operator moves in advance of any task feedback, with accurate initial judgment reducing the need for subsequent corrections.

The remote environment was made visually sparse by limiting content to only the stationary target markers and the plain taskboard background shown in Fig. 2. This removed obvious twist cues from the visual background that might otherwise have aided in the experimental judgment task by allowing subjects to estimate absolute rotations. Furthermore, by eliminating any subjective visual impression of an unsteady horizon, this reduced the potential for motion sickness. Interestingly, this sparse environment approximates the emptiness of space while on-orbit as well as the often murky, low contrast views encountered in underwater (and potentially for remote planetary) telepresence systems. This sparseness is also typical of VE simulations in which detail is reduced to lessen computational burden.
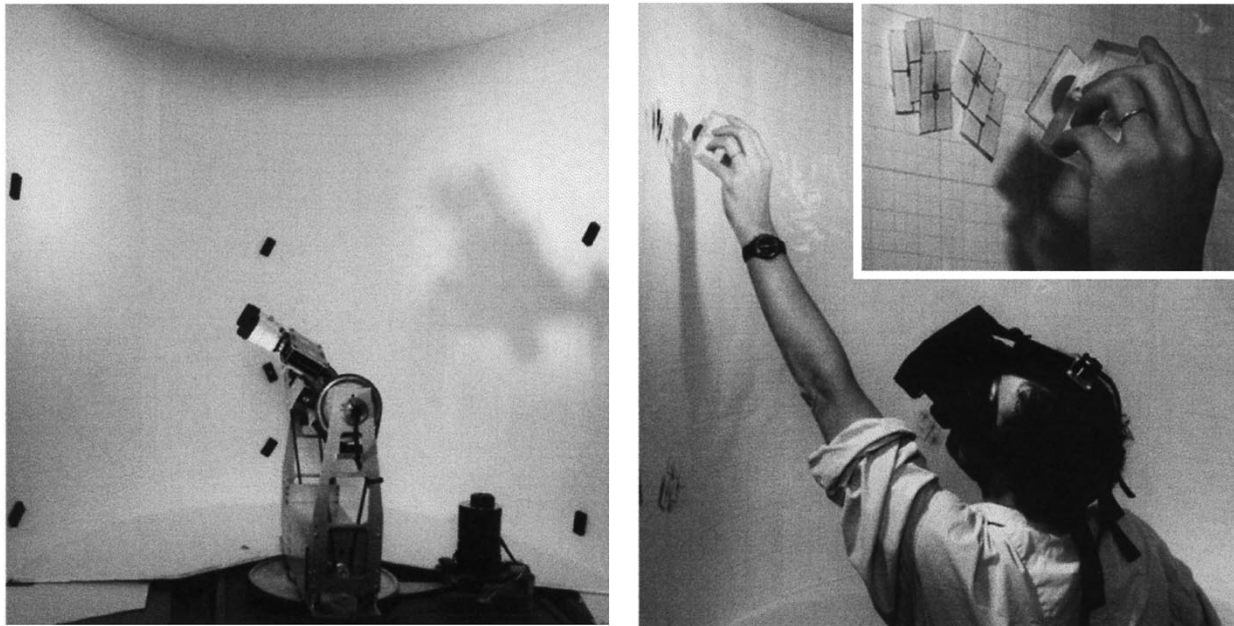
Fig. 2.　Remote station (left) and local station (right). Dispensing block and response tags (inset).

B. Preexperiment Setup and Alignment

At the start of the experiment, the subject's seating position was aligned so that head yaw (in the Fick system) corresponded approximately with the local taskboard's cylindrical axis. The subject was asked to maintain this posture in a standard backed chair during the experiment to minimize translation from sources other than those associated with neck biomechanics. The subject's body was not otherwise constrained. Software initialization of the camera platform followed the subject's indication that he was looking straight ahead and that his head felt level with respect to his internally perceived horizon. This head posture defined the subjective yaw-pitch-roll reference frame to which the spatial sensor output was rezeroed. Initialization was completed by moving each platform DOF to its true zero angle and setting this orientation in the servo control software to equal the rezeroed sensor frame. Thus, true level of the camera platform corresponded to subjective level in the HMD.

Following adjustment of binocular offset to allow fusion of the video images to the two eyes, a selection of fixed target marker positions was established for use throughout the entire experiment with that particular subject. The "reference" direction (approximating the primary position) was determined by having the subject gaze straight ahead—parallel to his internally perceived sagittal and horizontal planes—and guide the experimenter to place a target marker on the remote taskboard at the center of the monocular, circular reticle visible in the HMD.

The subject was instructed to use this reticle during the set up and actual experiment periods to align his gaze with the remote targets of interest. The reticle, which was inserted in the right camera lens system and therefore appeared fixed in that eye's LCD, subtended a visual angle of $5°$. Because the reticle was symmetric about the camera's optical axis, it did not provide orientation cues and thus offered no guidance to the subject on rotation (i.e., twist) about the gaze axis. The reticle served three purposes:

1) it forced subjects to use the same eye when aligning their gaze at the target;
2) it required them to look through the same region of the lens system when performing this task;

3) it ensured that they visually acquired the target through the center of the HMD and camera lenses, where optical distortions were at a minimum.

In principle, since the HMD was fitted snugly to the subject's head, the fixed reticle demanded uniform final eye-in-head gaze direction. Consequently, only head direction varied with final gaze alignment. Since eye movements (i.e., saccades) precede those of the head when naturally seeking out visual targets (e.g., [13]–[16]), the reticle was not expected to restrict initial eye-in-orbit motion, but only final direction when fixating.

After the reference target was placed, the subject was asked to yaw his head without moving his torso to a maximum azimuth on his left side and pitch upward to the highest attainable elevation, and then guide the experimenter to align a new marker within the reticle for the next target location on the remote taskboard. This process was repeated for maximum elevation at maximum azimuth on the right side and again for maximum elevation at the center azimuth above the reference target. Three more markers were placed at approximately the same three azimuths but at a downward elevation. The target markers, shown in the left panel of Fig. 2, were 47.6 by 22.2 mm rectangles that could be attached magnetically at any arbitrary position (i.e., direction) and angle (i.e., rotation) on the remote taskboard's inner surface.

After initial placement, all seven targets were adjusted as necessary to ensure that the subject, using the hand on the side of the target, could touch the local taskboard at the locations which he judged to correspond to the target positions on the remote board. The targets were then rotated arbitrarily and uniquely for each subject to "noncanonical" angles (i.e., not at $0°$, $30°$, $45°$, $60°$, $90°$, etc.) and left in place for the duration of the experiment.

The need for extreme azimuth-elevation combinations for the upper left and right targets followed from a pilot study in which more modest head pitch and yaw did not reveal noticeable platform roll related effects in either direction or orientation judgments. Glenn and Vilis's [9] observations further suggested that any roll effect would be greater where the twist angle mismatch between Listings law (or intermediate Fick-like head motion) and two DOF Fick camera platform kinematics is more pronounced—i.e., at large combined head yaw and pitch away from the primary position.

C. Experiment

At the beginning of the actual experiment, each subject was instructed to first visually locate specific target markers on the remote taskboard by moving his head to the target. Once the target was aligned with the reticle, the subject hand-placed a rectangular (76.2 × 25.4 mm cross section) dispenser block on the local taskboard (shown in Fig. 2), releasing a sequentially numbered response tag to match the perceived location and rotation of the remote target. Because of the tags' adhesive backing, response location and rotation could not be modified after initial placement on the taskboard. The subject used his left hand for targets on the left side and right hand for targets on the right, with no specific instruction given for center targets. After placing the response marker on the local task board, the subject was handed a dispenser loaded with a new tag and the next target was called out by the experimenter. The seven targets in each block were always called out in the same sequence: Center, Upper Left, Lower Right, Upper Center, Lower Left, Upper Right, and Lower Center (henceforth labeled CC, UL, LR, UC, LL, UR, and LC, respectively). This ordering separated targets widely enough on the remote taskboard such that consecutive pairs were not visible in the HMD at the same time, thereby eliminating relative targeting cues. To further reduce orientation cues, the subject was instructed to make judgments only after completing the gaze directed motion to the target once his head was no longer moving.

The principal factor of interest in this study was whether head-slaved roll was enabled in the camera platform. When enabled, the platform followed the full 3-D (yaw- pitch-roll) orientation of the subject's head. When disabled, roll was fixed in the Fick gimbal system such that the baseline joining the two camera axes remained parallel to the horizontal regardless of the subject's three DOF head orientation. The camera platform roll condition was switched after each block of seven targets was acquired. The subjects kept their eyes closed when the condition was switched and were never informed by the experimenter as to the platform's roll condition. In particular, subjects were monitored for any deliberate twisting motions that they might employ to test the camera platform's roll condition. No such motions were observed during the experiments.

Six right-handed male subjects, ages 23–47 years of age, participated in the full protocol. Half the subjects began with roll enabled; half with it disabled. Thus, the experiment design crossed target location with roll condition and nested the roll enabled/disabled sequence. The block of seven targets was repeated seven times under each roll condition for a total of 98 target acquisitions per subject. Each subject completed the entire experiment during one 45 min session. Explicit constraints were not imposed on the response time for individual targets. However, given the time to hand out response markers, to toggle the platform condition between blocks, and for short rests, individual targets were typically acquired within 20 s.

Because the targets were relatively small (2.0° × 4.5° in visual angle), subjects fixated within a narrow region during final target acquisition. Thus, the experiment task was designed to stimulate aiming of gaze, not the scanning of object boundaries or contours. The task was therefore intended to provide insight into how subjects evaluate gaze direction (and local rotation in those directions), and ultimately how they might incorporate this information into spatial awareness of their surroundings. Examples of local, telepresence, or VE applications that this experiment might represent include (but are not restricted to) control tower monitoring of 3-D air traffic and flight deck navigation. The use of manual pointing to indicate subject response may also contribute to understanding aspects of open-loop grappling or assembly in situations where excessive time delay or insufficient display hardware degrades haptic and visual feedback.

D. Data Processing

Raw target and response marker positions and rotations were measured by scale and protractor with respect to horizontal and vertical grid lines that permanently covered the inner surface of both taskboards. Because of red filters mounted on the camera lenses, the yellow lines and markings that made up the grid were invisible to the subjects. The scale and protractor had resolutions of 0.8 mm (1/32 in) and 0.5°.

Raw Cartesian position and rotation from the inner surface of the task boards were converted in stages into 3-D spherical coordinates. In the first stage, the CC target's raw horizontal (circumferential) and vertical coordinates were subtracted from the respective raw components for all that subject's response and target markers. This yielded the horizontal and vertical target and response *offset* relative to the individual subject's initially indicated "straight ahead" direction and had the effect of realigning all subjects' responses relative to the same common center direction. The offset of the twist response was determined by subtracting raw target from raw response rotations.

Systematic response variations over the course of the experiment were addressed by treating each seven-target block for changes in response bias. These changes could be caused by drifts in camera platform offset, subject seating offset, spatial sensor alignment with respect to the HMD, or HMD alignment relative to the subject's head, or by subject fatigue. Component means were calculated for each seven-target block of each subject's response offset. Next, respective block component means were subtracted from each of the seven responses making up that particular block. The average of that subject's seven target offsets was then added back to all his response offsets to yield the components of the *debiased* response. By accounting for the mean of each subject's seven targets, block bias subtraction is compensated for correction asymmetries caused by initial target placements with respect to CC—in effect basing the bias correction on the difference between individual responses and their respective targets.

The debiasing procedure had the desired effect of reducing response offset variation as measured from the complete experiment data set as well as from different subsets of targets, subjects, and roll condition. Averaged across all blocks for all subjects ($n = 14 \times 6 = 84$), debiasing shifted the vertical component of the response 8.5 ± 5.0 cm (mean $\pm$ stdev) upward (i.e., response offset was initially biased downward by this amount), the horizontal component 1.0 ± 3.0 cm to the right (i.e., initially biased left), and the twist by $8.0° \pm 7.1°$ clockwise (i.e., initially biased counter-clockwise). The directional shifts are equivalent on the 0.61 m radius taskboard to $7.9° \pm 4.6°$ in elevation and $0.9° \pm 2.8°$ in azimuth at target CC. When the two platform roll conditions were pooled separately, the bias corrections did not exhibited statistically significant linear trends versus block number (i.e., versus time).

Next, horizontal and vertical components of individual debiased responses and targets were converted into spherical azimuth and elevation directions (yaw and pitch in a Fick system) by simple trigonometry under the assumption that all pointing directions passed through the respective taskboard's axis of curvature at the height of target CC.

Finally, response *error* was computed for each judgment from the quaternion-based 3-D orientation transformation which uniquely maps target azimuth, elevation, and rotation into those of the response by a single angular twist (e.g., see [17]). This transformational approach produced less data scatter when compared with arithmetically subtracting individual response from target components.

IV. RESULTS

Mean response error directions and rotations were computed at each target location for the roll-enabled and roll-disabled camera platform conditions by arithmetic averaging of separate azimuth, elevation, and
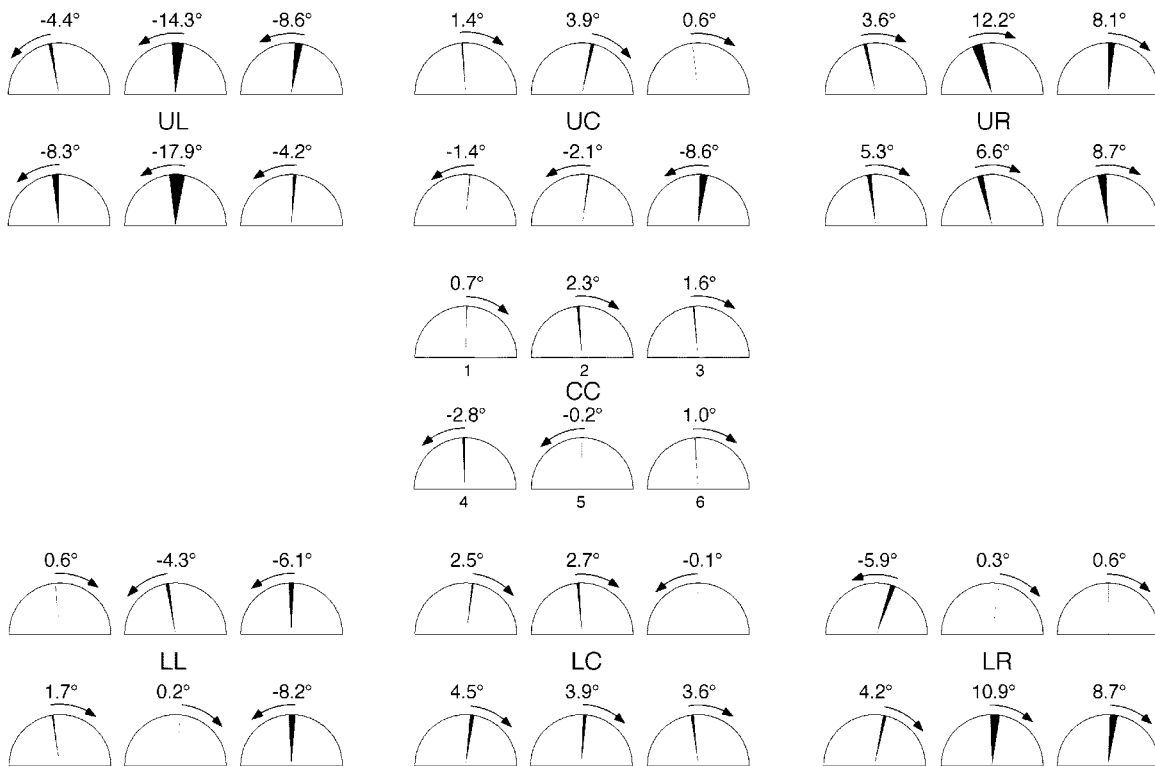
Fig. 3.   Average rotational components of response error for each subjects. Subject number $(1, \cdots, 6)$ shown at CC is the same for all target regions. The black segment in each case is bounded by the average rotational component of the response for the roll enabled (R) and disabled (NR) camera platform conditions. The angular size of the black segments and direction of the arrows represent the magnitude and the clockwise-counterclockwise (CW/CCW) sense of differences in rotation response error between the two conditions (NR minus R). The magnitude and direction of this difference is also indicated numerically above each icon.

rotation angles (i.e., sum of these individual error components derived from the error quaternion, divided by number of repetitions). Other methods for estimating mean errors, including singular value decomposition of summed rotation matrices of orientation errors [18] and normalized quaternion sums extended from [18], yielded nearly identical results. By averaging each of the three orientation error components separately, we were then able to employ conventional ANOVA for subsequent statistical analyzes.

A. Rotational Errors

The rotational component of mean response errors for each subject at the seven target locations is plotted in Fig. 3. The plot format shows the mean errors for the no-roll (NR) and roll (R) conditions as well as the difference between means for the two conditions. While the magnitudes of rotational error differences between the two conditions are typically small, they are more pronounced for most of the subjects at the UL and UR target locations. Furthermore, error difference signs indicate that each subject's mean response rotation was more counterclockwise for UL and more clockwise for UR targets when platform roll was disabled.

The separate ANOVA performed on the rotational component showed target location (TL) $(F_{6,24} = 3.744; p < 0.009)$ and the interaction between TL and roll condition (RC) $(F_{6,24} = 8.552; p < 0.001)$ were significant. No other statistically significant influences $(p < 0.05)$ of other factors, either alone or in combination, were observed. Due to the absence of significant roll sequence (RS) effects, there was no indication between subject groups of asymmetric transfer arising from the camera platform's starting RC.

The mean values of rotational components for each target region, averaged across all subjects are reported in Fig. 4 in the same format as
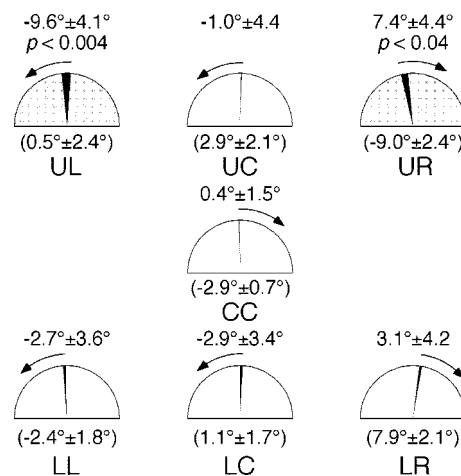

Fig. 4.   Response error rotational components (mean $\pm$ standard error) averaged across all subjects. Magnitude and CW/CCW sense of differences in rotation response error between the two camera conditions (NR minus R) is represented by the black segments' angular size and the arrows' direction, and is indicated numerically above each icon. Shading and associated prob-values indicate which differences are significant. Mean rotations (shown in parentheses as mean $\pm$ standard error) combining R and NR conditions are not significantly different from zero at any target location.

the previous plot. The mean combining both R and NR rotational errors for each target did not differ significantly from zero in any one region as tested by Scheffé contrast—critical $(p = 0.10)$ rotation magnitude of $9.5°$ for variance $s_w^2 = 88.322$ and $N = 12$ observations. For differences between roll and no-roll conditions (i.e., NR minus R), however, Scheffé contrasts between means were significant only for the coun-
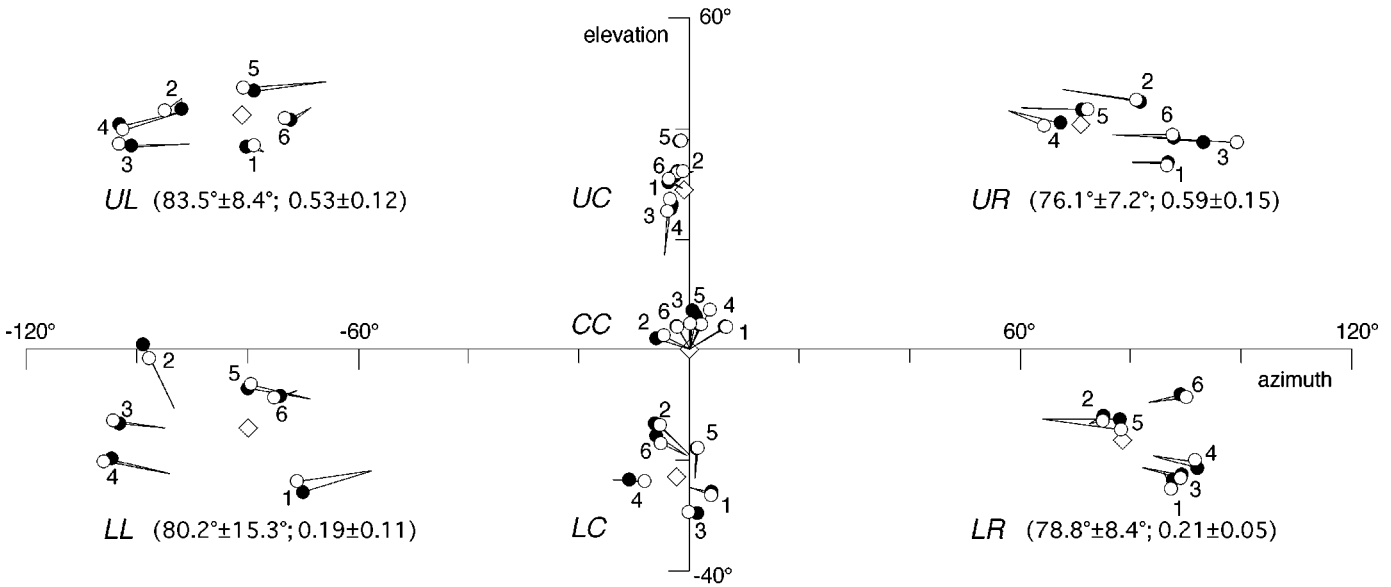
Fig. 5. Directional components of the targets and debiased responses averaged for each of six subjects (1, $\cdots$, 6) in each target region. Filled circles represent average responses with camera platform roll compensation enabled; unfilled circles the average with roll compensation disabled. Line segments join the average responses (circles) for the particular subject to the corresponding target location (unmarked end of each line segment), indicating the misjudgment direction and magnitude. The diamonds indicate the average target azimuths and elevations across all subjects for each region. Associated with labels UL, UR, LL, and LR are the respective target eccentricity (i.e., the great arc angular deflection combining azimuth and elevation relative to primary position CC) and elevation/azimuth ratio, averaged (mean ± stdev) across all subjects.

terclockwise value (i.e., NR more counterclockwise) at UL and clockwise value (i.e., NR more clockwise) at UR. The critical ($p = 0.10$) rotational difference was 6.4° for variance $s_w^2 = 10.141$ and $N = 6$ observations.

B. Directional Errors

Azimuth and elevation components of the mean response errors plotted in Fig. 5 show the tendency of all the subjects to overestimate azimuth for the large eccentricity targets (i.e., UL, LL, UR, LR), regardless of RC. Fig. 5 also indicates that the target marker locations in each of the six off-center regions differ from subject to subject, as would be expected given individual head motion and arm reach capabilities. Target CC always remains at the origin because of the offset adjustment procedure discussed above in Section III-D. Furthermore, all targets and responses lie between −40° and +60° in elevation, thereby ensuring that the camera platform and subjects' heads remained well away from the Fick gimbal kinematic singularities at ±90°.

The ANOVA conducted separately on the azimuth component of response error indicates a main effect ($F_{6,24} = 13.485$; $p < 0.001$) for TL. Neither RC nor RS alone, nor interactions between any of the factors had an effect on azimuth error ($p > 0.05$). Mean azimuth errors were significantly different from zero for UL ($p < 0.046$), LL ($p < 0.011$), and UR ($p < 0.002$), and nearly significant for LR ($p < 0.103$) by Scheffé contrasts for variance $s_w^2 = 42.634$ and $N = 12$ observations.

Similarly, an ANOVA on the elevation error component showed an effect ($F_{6,24} = 3.563$; $p < 0.011$) for TL, but none for RC, RS, or any interactions between factors. Again no evidence of asymmetric transfer due to RC order was detected in the direction results. Scheffé contrasts, however, did not show significant differences from zero for mean elevation errors at any individual TL—critical ($p = 0.10$) magnitude 4.6° for variance $s_w^2 = 21.199$ and $N = 12$ observations.

V. DISCUSSION

A. Rotational Component

Changes in the rotational component of subject response with RC, though small ($<10°$), were statistically significant, specifically at UL and UR, where target azimuths and elevations were largest. These targets correspond to the directions where Bolas and Fisher [5] observed operators to have most difficulty with a virtual Fick camera platform simulation head-slaved in only pan and tilt (i.e., yaw and pitch) DOF's. These are also the directions where two DOF Fick gimbal twist (i.e., roll) would be expected to have its greatest departure from the Listing system [9].

The implication for this rotational misjudgment is clearly task dependent. For example, if the objective is to align and assemble components in the absence of timely (or any) visual or other feedback, a 10° misjudgment can be an enormous impediment to successful task completion. Conversely, in 3-D air traffic visualization, threat detection depends on knowledge of target direction and probably would not require any assessment of rotation.

Platform roll condition and the rotational component of subject target judgment are related as follows. When the subject's head (with HMD and attached tracker) rolls by angle ϕ and platform roll is enabled, the slaved cameras also roll by ϕ. Consequently, the target image, which is fixed in the remote frame, rotates by $-\phi$ on the HMD's LCD panels. Because head-HMD rotation is exactly equal and opposite to image rotation on the LCD's, the subject sees no apparent target rotation. However, when platform roll is disabled, the target image cannot rotate on the LCD's and therefore appears to the subject as having moved along with his head in the local frame through angle ϕ. Accordingly, if the subject is unaware of the platform RC, the judgment difference between the no-roll and roll conditions (i.e., NR minus R) would have the same sign and magnitude as the subject's underlying head roll motion (i.e., $\phi - 0 = \phi$).

Models of gaze-directed head and eye orientation [7]–[9] which relate head roll angle to pointing direction can be used to predict our
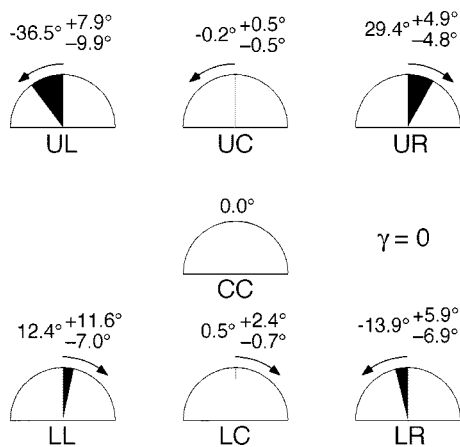
Fig. 6. Roll angle differences between two DOF Fick camera platform and subject head with Listing orientation. Each symbol represents the average (mean ± range) of difference angles ϕ_δ calculated from the six subjects' individual target placements.
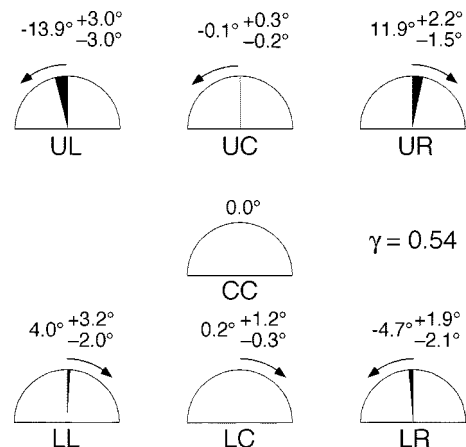

Fig. 7. Roll angle differences between two DOF Fick camera platform and subject head with Fick-like orientation. Each symbol represents the average (mean ± range) of difference angles ϕ_δ calculated from the six subjects' individual target placements.

subjects' differences in target rotation judgments between the two platform roll conditions. The roll angle expected for specific head pitch-yaw gaze directions toward our subjects' targets is calculated from the "gimbal score" formulation in the Appendix. The gimbal score, introduced by Glenn and Vilis [9], interrelates the directional components of the instantaneous axis of rotation (i.e., the quaternion or Euler axis) for hypothetical head-pointing mechanisms. For example, when the roll DOF is disabled and leveled, the Fick camera platform mechanism used in these studies has a gimbal score of $\gamma = -1$ and, therefore, a roll angle of $\phi = \phi_F = 0$. Straumann *et al.* [8] and Tweed and Vilis [7] interpreted data for the head in space as following Listing's law, corresponding to a gimbal score of $\gamma = 0$ and producing a roll angle $\phi = \phi_L$ when expressed in the Fick sequence of angular coordinates discussed in the Appendix. Therefore, when Listing head orientation is tracked by a Fick gimbal capable of only two DOF's, the mismatch between head and camera platform roll is given by: $\phi_\delta = \phi_L - \phi_F = \phi_L$. As proposed above, this mismatch between head and camera platform angles corresponds to the difference in roll judgment between platform conditions.

Fig. 5 shows the target directions for individual subjects that were employed in estimating judgment differences. The oblique targets (UL, UR, LL, and LR) all have large eccentricities, on average commensurate with the respective 70 and 90° target eccentricity of [7] and [9]. In the prior studies [7]–[9], however, eye motion relative to the head was not restricted while targets were set by the experimenters to have equal elevation and azimuth—i.e., vertical-to-horizontal ratios (v/h) with unit magnitude. In our experiment, final eye-in-head direction was fixed by the reticle in the right LCD while subjects selected their own target locations based on maximum individual head orientation and arm reach capabilities. Since the relative contributions to overall gaze direction of head versus eye motion are greater for azimuth than elevation [7]–[9], v/h measured for head deflection in [9] was reduced, respectively, to 0.54 ± 0.01 and 0.50 ± 0.11 for 70 and 90° target eccentricities, consistent with the placement of our UL and UR targets. LL and LR targets had smaller v/h magnitudes because the camera platforms base and resulting taskboard design precluded more negative (i.e., downward) target elevations.

Fig. 6 summarizes the rotational judgment differences predicted from the yaw and pitch direction to each subject's targets (i.e., $\phi_\delta = \phi_L$) for head orientations constrained by Listing's law (i.e., $\gamma = 0$). While the signs predicted by the Listing model in Fig. 6 agree with the measured differences for the upper corner targets of Fig. 4, the model's magnitudes are much larger. Furthermore, the Listing model

indicates that relatively large head roll and consequent rotational judgment differences should occur at the lower corners—something not evident in the responses shown by Fig. 4. Glenn and Vilis [9] observed, however, that when head v/h ratios and target eccentricities extend into ranges similar to our experiments, head-in-space orientation exhibited Fick gimbal-*like* rather than Listing characteristics, with an intermediate gimbal score of $\gamma = -0.54 \pm 0.19$ (mean ± standard deviation). Using this "Fick-like" γ, the mismatch between head and camera platform roll in Fig. 7 maintains the same sign but is now within the 6.4° uncertainty (critical bound $p = 0.10$) of the judgment differences in Fig. 4.

Eye-in-orbit roll (torsion about the line of sight) was not included as part of the head orientation model because it was not expected to affect target judgment between platform conditions. One component of eye-in-orbit torsion, so-called "false" torsion, is simply an artifact of the coordinates (e.g., Fick system) chosen to express the 3-D orientation caused by two-dimensional Listings law eye-in-orbit gaze deflection [6, p. 185]. This torsional component was essentially zero in our experiment, because the reticle restricted final eye-in-orbit to near its primary position at which twist does not occur, irrespective of coordinate convention. Another component is *counter*torsion of the eyes (cycloversion), generated by the utricular and neck muscle reflexes, in opposition to head roll [6, pp. 383–384]. This component, however, results in no more than ~2° of eye-in-orbit torsion [19], even for the peak head roll angles estimated to have occurred at the upper corners. Furthermore, because of identical targets and consequent induced head roll, eye torsion from these sources should not vary across platform conditions.

B. Directional Components

Subjects overestimated the azimuth of large eccentricity targets, regardless of RC, with statistical significance at UL, LL, and UR, and near significance at LL. Though we detected an effect of target location on elevation response error by ANOVA, *post hoc* contrasts did not reveal the elevation misjudgments to be significant at any specific target location.

Misrepresentations of perceived direction have been reported in "open loop" head pointing studies where subjects judged the combined azimuth and elevation of icons in an exocentrically arranged perspective display [20], yaw angle in exocentrically presented avionics displays [21], yaw angle of an imagined, internally-referenced horizontal clock [21], and the azimuth of localized aural stimuli [22].

Similarly, related errors have also been observed for hand pointing to visually acquired azimuth target [14], [16], [23]. A model-based interpretation of the direction misjudgments in this and the prior studies in [20] and [21] will be reported elsewhere.

C. Implications for VE and Telepresence Systems

The addition of the roll DOF to the Fick gimbal camera platform in our experiment had a statistically significant effect on subjects' ability to judge the rotation of simple rectangular objects in a quasi-static task, specifically for the targets which combined large head pitch and yaw excursions. This effect, though significant, was relatively small, with average rotation judgment difference magnitudes between platform conditions of 7.4 and 9.6°, respectively, at UL and UR. Target elevations in our experiment, however, were limited upward to the highest point on the local taskboard that could be touched by the subject and downward by the structure of the camera platform and taskboard. Furthermore, our tests were conducted with the camera platform mechanism's primary position (straight ahead direction) adjusted to coincide with that of the subject. Had greater elevation or depression (i.e., negative elevation) magnitudes been possible, or had the camera platform's base been pitched upward or downward such that it was vertically tilted with respect to the subjects' primary position, much larger rotation judgment differences between platform conditions would be expected. This expectation is based on the Fick-like head model's greater departure from the orientation of the two DOF Fick gimbal at larger elevations (depressions).

Additionally, at the most extreme elevations, our platform's controllability and pointing accuracy would suffer because of the kinematic singularities intrinsic to its Fick gimbal architecture—both upward and downward along its vertical axis. Other two DOF camera platform configurations would have different efficacy in tracking subject head orientation. For example, the Helmholtz gimbal shown in Fig. 1—in which the pitch DOF supports yaw (opposite of the Fick configuration)—has singularities at either side of its horizontal axis. Because a two DOF Helmholtz platform would have gimbal score $\gamma = 1$, its roll angle would deviate further than a Fick mechanism ($\gamma = -1$) from Fick-like ($\gamma = -0.54$) head motion as gaze is directed away from the primary position.

In some applications (e.g., head-tracked cockpit simulators or see-through HMD's for combat aircraft), operators must attend frequently to "over-the shoulder" events that necessitate oblique, high-eccentricity gaze direction maneuvers. The implication of our results is that, except for these extreme and often uncomfortable head yaw-pitch combinations, the inclusion of a third roll DOF might not sufficiently alter the ability to judge the direction and rotation of stationary objects in the visual scene. In other situations, it might be advantageous to avoid circumstances leading to these extreme head-neck orientation combinations and the consequent need for the additional roll DOF. For instance, workstations for some immersing environments could be designed to have the operator sit in a pivoting chair or stand such that view azimuth is changed by yaw of the whole body while pitch by the neck directs gaze elevation. Thus, the increased expense of adding this third DOF to head-slaved camera platform mechanisms, movable boom-type viewers, or computer generated simulations may sometime not be warranted by the limited benefits to spatial situation awareness.

The sign and magnitude of rotational judgment differences induced by roll condition are predictable from the mismatch between the orientational kinematics of the human head and a Fick gimbal camera platform restricted to two DOF's. However, once roll is enabled, the general three DOF capability of our platform allows any arbitrary orientation scheme (provided we remain sufficiently distant from hardware singularities), including Listing and Fick-like strategies, to direct and twist the cameras. Because Listing and Fick-like kinematics have

only two independent angular coordinates (ψ and θ) as shown in the Appendix, it is plausible to design camera platform and viewer mechanisms that do not require three independently actuated DOF's to provide accurate tracking of human head orientation. A mechanism that has a gimbal score, γ, which matches the yaw-pitch-roll capabilities of the human head could reduce rotational misjudgments, yet accomplish this with only two actuators driving the two independent DOF's, ψ and θ.

Open-loop response errors, such as those reported here, however, can be diminished by improving visual or kinesthetic feedback from the virtual or remote environment pointing task (e.g., [23]). This feedback may aid in correcting rotational misjudgments induced by inaccurate tracking and rendering of the operator's head roll angle. Specifically, the operator's manual performance could benefit from the "visual reafference" offered by partial or sustained vision of his own hand [13], [24], as well as contextual information [24] from the remote or virtual environment that were not available in our visually sparse testing protocol.

APPENDIX

For the Fick gimbal mechanism of Fig. 1 in which the yaw (ψ) axis supports pitch (θ) followed by a roll (ϕ) axis, the three rotational displacements (i.e., Euler angles) that contribute to the 3-D orientation of the endpoint can be represented by the quaternions

$$\boldsymbol{q}_\psi = \begin{bmatrix} C_{\psi/2} \\ 0 \\ 0 \\ S_{\psi/2} \end{bmatrix} \boldsymbol{q}_\theta = \begin{bmatrix} C_{\theta/2} \\ 0 \\ S_{\theta/2} \\ 0 \end{bmatrix} \boldsymbol{q}_\phi = \begin{bmatrix} C_{\varphi/2} \\ S_{\varphi/2} \\ 0 \\ 0 \end{bmatrix} \quad \text{(A.1)}$$

where $c_{\psi/2} = \cos(\psi/2)$ and $s_{\psi/2} = \sin(\psi/2)$, etc. The single quaternion describing the sequence of rotations from the fixed base outward to the gimbal endpoint (e.g., camera on the platform or human head on a torso) is developed through the quaternion multiplication $\boldsymbol{q}_{\phi\theta\psi} = \boldsymbol{q}_\phi \circ \boldsymbol{q}_\theta \circ \boldsymbol{q}_\psi$ (e.g., [17]). The 3-D orientation of the gimbal endpoint, expressed in terms of the fixed base coordinate system, given by the quaternion inverse $\boldsymbol{q}_{\phi\theta\psi}^{-1} = \boldsymbol{q}_\psi^{-1} \circ \boldsymbol{q}_\theta^{-1} \circ \boldsymbol{q}_\phi^{-1}$, is

$$\boldsymbol{q}_{\phi\theta\psi}^{-1} = \begin{bmatrix} S_{\psi/2}S_{\theta/2}S_{\phi/2} - C_{\psi/2}C_{\theta/2}C_{\phi/2} \\ C_{\psi/2}C_{\theta/2}S_{\phi/2} + S_{\psi/2}S_{\theta/2}C_{\phi/2} \\ C_{\psi/2}S_{\theta/2}C_{\phi/2} - S_{\psi/2}C_{\theta/2}S_{\phi/2} \\ S_{\psi/2}C_{\theta/2}C_{\phi/2} + C_{\psi/2}S_{\theta/2}S_{\phi/2} \end{bmatrix}. \quad \text{(A.2)}$$

Glenn and Vilis [9] defined the gimbal score, γ, which best fits the quaternion torsional component, q_1, to the corresponding vertical and horizontal components, q_2 and q_3, for a set of experimental head or eye orientation data according to the relation

$$q_1 = \gamma(q_2 q_3)/q_0 \quad \text{(A.3)}$$

q_0 and (q_1, q_2, q_3), respectively, the scalar and vector components of the gimbal system quaternion described by $\boldsymbol{q}_{\phi\theta\psi}^{-1}$, are the first through fourth rows of (A.2). The quaternion's vector components define the orientation of an angular deflection's rotation axis—i.e., its Euler axis. The vertical, horizontal, and torsional component labels used in [9] correspond respectively to the yaw, pitch, and roll axes of the Fick system when in the primary (straight ahead) position.

After rearranging (A.3) and substituting for (A.2), the gimbal score can be expressed as

$$\gamma = \\ -\left[\frac{\sin \psi \sin \theta \cos \phi + 2 \sin \phi \left(\cos^2 \psi/2 - \sin^2 \theta/2\right)}{\sin \psi \sin \theta \cos \phi + 2 \sin \phi \left(\sin^2 \theta/2 - \sin^2 \psi/2\right)} \right]. \quad \text{(A.4)}$$

Alternatively, when the gimbal score γ is known, (A.4) can be inverted to solve for roll, ϕ, as a function of yaw-pitch angles, ψ and θ, according to

$$\phi = -\tan^{-1}\left[\frac{0.5(1+\gamma)\sin\psi\sin\theta}{(1+\gamma)\sin^2(\psi/2)+(1-\gamma)\sin^2(\theta/2)-1}\right]. \tag{A.5}$$

The two DOF Fick gimbal, has gimbal score $\gamma = -1$ for which (A.5) yields, as expected, $\phi = 0$. For a Listing system, $\gamma = 0$, which from (A.3) causes $q_1 = 0$—the constraint of Listing's law, that the rotation axis (i.e., quaternion vector) must lie in the (q_2, q_3) plane, perpendicular to the primary position (i.e., perpendicular to the q_1 component direction). Furthermore, even though a gimbal can have three coordinates to describe its orientation, only two of these are independent in the Listing system. The intermediate configuration, $-1 < \gamma \leq 0$, is termed Fick-like [9]. The remaining regime of possible scores in (A.3), $0 < \gamma \leq 1$, comprising Helmholtz (Fig. 1) and Helmholtz-like gimbals, is not considered in this derivation.

ACKNOWLEDGMENT

REFERENCES

[1] T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, MA: MIT Press, 1992.
[2] C. P. Comeau and J. S. Bryan, "Headsight television system provides remote surveillance," *Electronics*, pp. 86–90, Nov. 10, 1961.
[3] S. Tachi and H. Arai, "Study of tele-existence (II) three dimensional color display with sensation of presence," in *Proc. Int. Conf. Advanced Robotics*, Tokyo, 1985, pp. 345–352.
[4] S. Tachi, H. Arai, and T. Maeda, "Development of anthropomorphic tele-existence slave robot," in *Proc. Int. Conf. Advanced Mechatronics*, Tokyo, 1989, pp. 385–390.
[5] M. T. Bolas and S. S. Fisher, "Head-coupled remote stereoscopic camera system for telepresence applications," in *Stereoscopic Displays and Applications*. Bellingham, WA: SPIE, 1990, vol. 1256, pp. 113–123.
[6] I. P. Howard, *Human Visual Orientation*. New York: Wiley, 1982.
[7] D. Tweed and T. Vilis, "Listing's law for gaze-directing head movements," in *The Head-Neck Sensory Motor System*, A. Berthoz, W. Graf, and P. P. Vidal, Eds. London, U.K.: Oxford Univ. Press, 1992, pp. 387–391.
[8] D. Straumann, T. Haslwanter, M. C. Hepp-Raymond, and K. Hepp, "Listing's law for eye, head, and arm movements and their synergistic control," *Exp. Brain Res.*, vol. 86, pp. 209–215, 1991.
[9] B. Glenn and T. Vilis, "Violations of Listing's law after large eye and head gaze shifts," *J. Neurophysiology*, vol. 68, pp. 309–318, 1992.
[10] B. D. Adelstein and S. R. Ellis, "Effect of head-slaved visual image roll on spatial situation awareness," in *Proc. 37th Annu. Meeting Human Factors Ergonomics Soc.*, 1993, pp. 1350–1354.
[11] B. D. Adelstein, E. R. Johnston, and S. R. Ellis, "A testbed for characterizing dynamic response of virtual environment spatial sensors," in *Proc. 5th Ann. Symp. User Interface Software and Technology*. New York: ACM Press, 1992, pp. 15–22.
[12] W. Robinett and J. P. Rolland, "A computational model for the stereoscopic optics of a head mounted display," *Presence*, vol. 1, pp. 45–63, 1992.
[13] Y. Rossetti, B. Tadary, and C. Prablanc, "Optimal contributions of head and eye positions to spatial accuracy in man tested by visually directed pointing," *Exp. Brain Res.*, vol. 97, pp. 487–496, 1994.
[14] T. Uemura, Y. Arai, and C. Shimazaki, "Eye-head coordination during lateral gaze in normal subjects," *Acta Otolaryngol.*, vol. 90, pp. 191–198, 1980.
[15] W. H. Zangemeister and L. Stark, "Gaze types: Interactions of eye and head movements in gaze," *Exp. Neurol.*, vol. 77, pp. 563–577, 1982.
[16] B. Biguer, C. Prablanc, and M. Jeannerod, "The contribution of coordinated eye and head movements in hand pointing accuracy," *Exp. Brain Res.*, vol. 55, pp. 462–469, 1984.
[17] K. Shoemake, "Animating rotation with quaternion curves," *Comput. Graph.*, vol. 19, no. 3, pp. 245–254, 1985.
[18] W. D. Curtis, A. L. Janin, and K. Zikan, "A note on averaging rotations," in *Proc. IEEE Virtual Reality Annu. Int. Symp.*, 1993, pp. 377–385.
[19] H. Schöne, "Über den Einfluss der Schwerkraft auf die Augenrollung und auf die Wahrnehmung der Lage in Raum," *Z. vergl. Physiol.*, vol. 46, pp. 57–87, 1962.
[20] S. R. Ellis, S. Smith, and S. Hacisalihzade, "Visual direction as a metric of virtual space," in *Proc. 33rd Annu. Meeting of the Human Factors Soc.*, 1989, pp. 1392–1395.
[21] N. S. Dorighi, A. J. Grunwald, and S. R. Ellis, "Perspective format for a primary flight display and its effect on pilot situation awareness," in *Proc. 37th Annu. Meeting Human Factors Ergonomics Soc.*, 1993, pp. 88–92.
[22] D. R. Perrott, H. Ambarsoom, and J. Tucker, "Changes in head position as a measure of auditory localization performance: Auditory psychomotor coordination under monaural and binaural listening conditions," *J. Acoust. Soc. Amer.*, vol. 82, pp. 1637–1645, 1987.
[23] J. B. de Graaf, A. C. Sittig, and J. J. Denier van der Gon, "Misdirections in slow goal-directed arm movements and pointer-setting tasks," *Exp. Brain Res.*, vol. 84, pp. 434–438, 1991.
[24] J. L. Velay and D. Beaubaton, "Influence of visual context on pointing movement accuracy," *Cah. Psychol. Cogn.*, vol. 6, pp. 447–456, 1986.

Coordinated Planning and Control of Automated Assembly Manufacturing

Julius S. Gyorfi and Chi-Haur Wu

Abstract—A general, hierarchical method of planning coordinated motions among multiple objects in a dynamic environment is developed. This model consists of a two-phase approach: planning a global path for each object and then locally optimizing along that global path. The applicability of this model to surface-mount manufacturing systems is studied in different scenarios. By changing the priorities of the machines, different results may be obtained. By comparing these results, the best one may then be determined. Simulation results using the two-phase algorithm show an improvement over other planning and control methods.

Index Terms—Aseembly manufacturing, part placement, planning, surface-mount manufacturing.

I. INTRODUCTION

Efficient planning and control is a difficult problem for automated assembly systems. Component constraints, product specifications, and workspace limitations can combine to create a very complex set of motion constraints for the individual machines. In systems that contain many controllable machines, the total cycle time and throughput can be improved by coordinating the machines' motions. Coordinated motion allows multiple machines to perform their tasks in parallel. The performance improvement that can be realized is dependent upon the efficiency of the coordination.

The problem of motion planning may be divided into two stages: 1) navigation and 2) guidance [17]. In general, navigation is the problem