C-arm Calibration - is it really Necessary?

Ameet Kumar Jain[a], Michael An[b], Nicha Chitphakdithai[b], Gouthami Chintalapani[a], and Gabor Fichtinger[a,c]

[a]Department of Computer Science, Johns Hopkins University
[b]Department of Bio-medical Engineering, Johns Hopkins University
[c]Department of Radiology, Johns Hopkins University

ABSTRACT

C-arm fluoroscopy is modelled as a perspective projection, the parameters of which are estimated through a calibration procedure. It has been universally accepted that precise intra-procedural calibration is a prerequisite for accurate quantitative C-arm fluoroscopy guidance. Calibration, however, significantly adds to system complexity, which is a major impediment to clinical practice. We challenge the status quo by questioning the assumption that precise intra-procedural C-arm calibration is really necessary. Using our theoretical framework, we derive upper bounds on the effect of mis-calibration on various algorithms like C-arm tracking, 3D reconstruction and surgical guidance in virtual fluoroscopy - some of the most common techniques in intra-operative fluoroscopic guidance. To derive bounds as a function of mis-calibration, we model the error using an affine transform. This is fairly intuitive, since small amounts of mis-calibration result in predictably linear transformation of the reconstruction space. Experiments indicate the validity of this approximation even for 50 mm mis-calibrations.

The problem is twofold: (a) C-arm intrinsic calibration; and (b) C-arm distortion correction. Using our theoretical and experimental analysis on mis-calibrated C-arms, we propose a hypothesis that indicates that intrinsic calibration might not be required for a family for surgical procedures. The framework also makes suggestions to the current work flow so as to further minimize the effect of any inherent mis-calibration. To address the problem of pose dependant distortion correction, we propose the use of a framework that can statistically study the maximum variation in distortion near a certain pose and then intra-operatively use an average correction. These theoretical derivations and experimental results make a strong case for the use of mis-calibrated C-arms, obviating the cumbersome intra-operative calibration process, potentially boosting clinical applicability of quantitative fluoroscopy in many procedures. We also prove this hypothesis experimentally.

Keywords: C-arm Calibration, 3D fluoroscopic guidance.

1. INTRODUCTION

C-arm fluoroscopy is ubiquitous in general surgery, due to its real-time nature, versatility, and low cost. At the same time, quantitative fluoroscopy has not found a *large scale clinical acceptance*, because of some inherent technical difficulties. It needs to solve four major problems: (1) C-arm image distortion; (2) Calibration of model parameters; (3) Pose recovery or tracking when multiple images are taken; and (4) Registration to other imaging modalities. Quantitative fluoroscopy could enable a significant improvement in the current clinical practice, the prominent works tackling some of the above problems.[1,2]

If is known that both image distortion[3] and intrinsic calibration[4,5] may vary significantly with pose. The intrinsic parameters themselves can be classified into two broad categories - pixel size and location of X-ray source (focal spot) with respect to the image intensifier. The pixel size of a C-arm is usually constant throughout the life of the C-arm, depending on the actual electronic hardware. The location of the X-ray source on the other hand varies from pose to pose due to the flexibility in the mechanical design and the large weight of the X-ray source.

Honorable Mention Poster Award. Send correspondence to {jain, gabor}@cs.jhu.edu

Medical Imaging 2007: Visualization and Image-Guided Procedures, edited by Kevin R. Cleary, Michael I. Miga,
Proc. of SPIE Vol. 6509, 65092U, (2007) · 1605-7422/07/$18 · doi: 10.1117/12.712301

Image distortion usually has a consequential contribution to reconstruction error and needs to be compensated. Thus the additional cost of a full online calibration is not substantial. Recently developed advanced intensifier tubes allow for lesser distortion, while modern flat panel detectors obviate distortion correction altogether. This fact brings up the question whether we need to calibrate the C-arm fully at each pose. The question also leads to the broader issue, that even if it is not pose dependent, how accurate does calibration need to be. In spite of the importance of calibration in C-arm fluoroscopy, as far as the authors are aware, there has been no prior work that conducts this kind of analysis. The vision community has a similar problem[6, 7] when cameras are used for visual serving of robots. In essence, this paper addresses the issue of sensitivity of 3D fluoroscopy to mis-calibration.

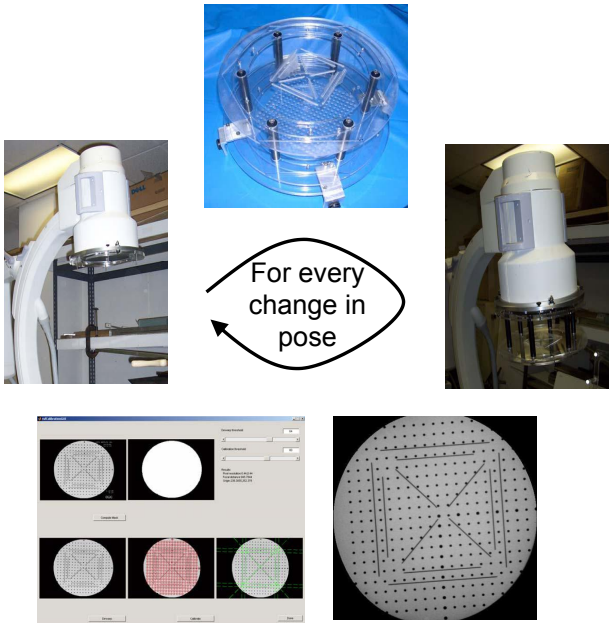


Figure 1. The C-arm intrinsic calibration is pose dependant. Hence conventional methods for fluoroscopic guidance need to calibrate the C-arm at every pose, a significant liability for intra-operative navigation. Traditionally, a calibration phantom is attached to the image intensifier, an X-ray image is taken and sent to a calibration module. The phantom is then removed, making the C-arm ready for use with the patient. This is repeated for every desired pose.

In quantitative C-arm fluoroscopy, we typically need to measure the spatial transformation between two objects, such as a vertebra and a bone drill, as compared to the transformation between an object and the C-arm itself. Thus the central intuition of this paper is that *while an incorrect calibration gives erroneous estimates for the absolute transformations, nevertheless it still provides acceptable relative estimates.* The consequence of this conjecture is potentially far reaching, as it can turn fluoroscopy to an affordable quantitative measurement tool in a large family of procedures. It should however be noted that we do not claim that calibration would always be unnecessary, since there are some applications that require accurate absolute estimates, as can be observed in cone beam computed tomography.[5] The decision should always be made case by case, experimentally.

The driving application for the proposed hypothesis is virtual fluoroscopy (VF), where synthetically created X-ray images in conjunction with image-guidance aids the surgeon in accurately targeting the anatomy.[8] Furthermore, it significantly decreases the amount of radiation dose imparted to the surgeon, which has been typically very high in these family of procedures. In this paper, we build a mathematical framework to formally address this issue and lend credit to the intuition that a loose estimate of the C-arm parameters might suffice. Furthermore, our theoretical derivations also provide us with minor modifications to the current VF procedures,

so as to minimize the effect of any inherent mis-calibration. In essence, we prove in theory and demonstrate experimentally that VF is feasible with an un-calibrated C-arm.

2. MATHEMATICAL FRAMEWORK

C-arm Imaging: Geometric aspects of fluoroscopic imaging can be modelled as a perspective transformation with five parameters - focal length, image origin and pixel size (Figure 2). The transformation formula is given in equation (1), (2); where F_O, F_X, F_I are the coordinate frames of an object, X-ray source and the image, respectively; P is a 3D point; P_O are the homogenous coordinates of P in F_O; p is the projection of P on the image plane; p_I are the homogenous coordinates (in pixels) of p in frame F_I; ${}^X F_O$ is the 4x4 rigid transformation matrix that transforms a point in F_O to F_X (also called pose); ${}^I F_X$ is the 3x4 perspective projection matrix; f is the focal length; $O = (o_x, o_y)$ is the projection of the X-ray source on the image plane (later referred to as the origin); s_x and s_y are the pixel sizes along the X and Y axis of the image.

$$p_I = {}^I F_X \;\; {}^X F_F \;\; P_F \tag{1}$$

$${}^I F_X = \begin{bmatrix} -f/s_x & 0 & o_x & 0 \\ 0 & -f/s_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad {}^X F_O = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_1 \\ r_{21} & r_{22} & r_{23} & T_2 \\ r_{31} & r_{32} & r_{33} & T_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

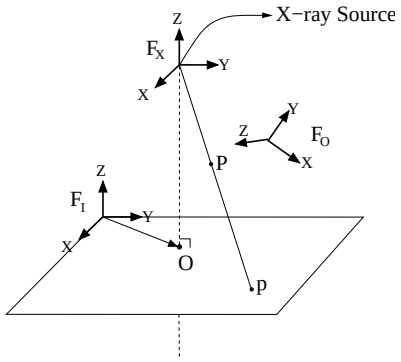


Figure 2. Projective geometry and notations for fluoroscopic imaging.

There are a total of five independent parameters that need to be evaluated by the calibration procedure - the pixel sizes (two) and the focal spot (three). Note that owing to the mathematics of projective imaging, these intrinsic parameters can only be estimated up to an arbitrary scale (which determines the physical size of the C-arm). Hence instead of the pixel sizes, the aspect ratio (ratio of pixel sizes) is typically used. Since the aspect ratio remains unchanged throughout the life of the C-arm, online-calibration essentially reduces just to the location of the focal spot. The proposed theoretical framework can be used to study sensitivity due to any of the five parameters, nevertheless, we limit ourselves only to that of the focal spot.

2.1. Model for Reconstruction Space Transformation

As illustrated in Figure 8, let A & B (with reference frames F_A & F_B) be the two objects being imaged. The assumptions are: (i) ${}^I F_A$,${}^I F_B$ can be computed from the images; (ii) A & B are *not large* in comparison to the focal length; (iii) F_A and F_B are *close by*; and (iv) the quantity of interest is ${}^A F_B = ({}^I F_A)^{-1} \; {}^I F_B$. Let $\bar{f}_1$ be the true focal spot and $\bar{f}_2 = (\bar{f}_1 + \bar{D})$ be the mis-calibrated estimate. We claim that even though the absolute locations of the objects are off, their relative transformation might still be accurate.

$$[{}^I F_A(\bar{f}_1)]^{-1} \;\; {}^I F_B(\bar{f}_1) \quad \sim \quad [{}^I F_A(\bar{f}_2)]^{-1} \;\; {}^I F_B(\bar{f}_2) \tag{3}$$

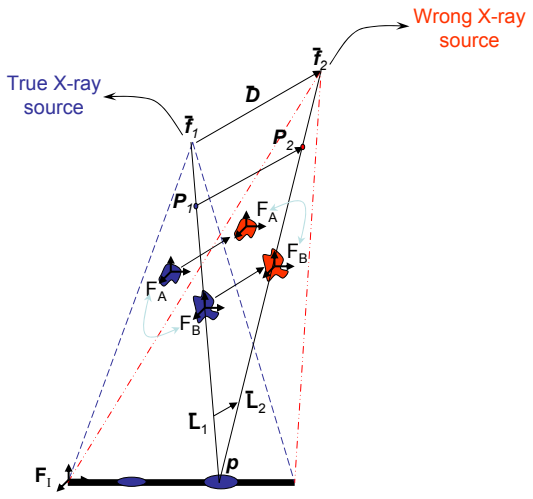


Figure 3. Mis-calibration shifts all reconstructed objects under an affine transformation.

A transformation is needed that can take the absolute location of an object reconstructed with calibration $\bar{f}_1$, and compute its corresponding location with calibration $\bar{f}_2$. We claim that the simplest transformation will be a linear affine model $\mathcal{T}$. The intuition derives from the observation that the image plane is the same in both reconstruction spaces. Thus if P_1 (not in homogenous coordinates) projects to a point p on the image, then it is constrained to be on line $\bar{L}_1$ in the $\bar{f}_1$-space and on $\bar{L}_2$ in $\bar{f}_2$-space. Thus we seek a continuous invertible transformation that projects $\bar{L}_1$ to $\bar{L}_2$. By incorporating the above constraints, $\mathcal{T}$ can be evaluated to be,

$$P_2 \;=\; \mathcal{T}\cdot P_1 \;=\; \begin{bmatrix} 1 & 0 & D_x/f_{1z} \\ 0 & 1 & D_y/f_{1z} \\ 0 & 0 & 1+(D_z/f_{1z}) \end{bmatrix} \cdot P_1 \;=\; P_1 + (d\cdot Z/f_{1z})\hat{D} \tag{4}$$

where with respect to (wrt) F_I, $\bar{D} = (D_x, D_y, D_z)$; $d = \|\bar{D}\|_2$; $\hat{D} = \bar{D}/d$; $\bar{f}_1 = (f_{1x}, f_{1y}, f_{1z})$; and $P_1 = (X, Y, Z)$. Each point is effectively translated in direction $\hat{D}$ by an amount proportional to its distance from the image. Experiments measuring the correctness of this affine model are available in Section 3. Thus to study sensitivity, it is sufficient to study the properties of $\mathcal{T}$.

Changes in Length and Scale $\mathcal{T}$ preserves the scale along the x, y-axes, but scales the space along the z-axis. Let $P_1(X_1, Y_1, Z_1)$ & $P_2(X_2, Y_2, Z_2)$ be any two points (not necessarily close to each other) in the $\bar{f}_1$-space at a distance of l_1. $\mathcal{T}$ maps them to points Q_1 & Q_2 in the $\bar{f}_2$-space at a distance of l_2 (Figure 4 (a)). It can be shown that

$$l_2^2 = l_1^2 + \frac{d^2}{f_{1z}^2}(Z_1 - Z_2)^2 + 2\frac{(Z_1 - Z_2)}{f_{1z}}\bar{D}\cdot(P_1 - P_2) \tag{5}$$

$$\|l_2 - l_1\| \;\leq\; \frac{d}{f_{1z}}|Z_1 - Z_2| \tag{6}$$

$$\Longrightarrow \quad l_1[1 - \frac{d}{f_{1z}}] \;\leq\; l_2 \;\leq\; l_1[1 + \frac{d}{f_{1z}}] \tag{7}$$

It directly follows from Equation (6) that $\mathcal{T}$ does not alter the length significantly. As an example, a 10 mm calibration error would affect the length of a 30 mm thoracic pedicle screw at an angle of 45^o by less than 0.2 mm (focal length $\sim$ 1 m), which is significantly less than the error from other sources. Thus F_A, F_B will not change their relative translation by a factor more than that specified by Equation (6).

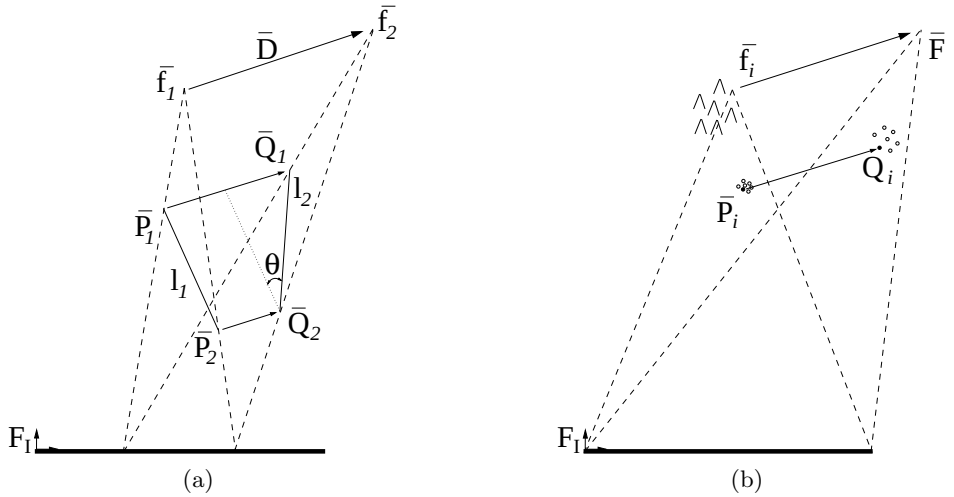


Figure 4. (a) Mis-calibration rotates and scales a straight line segment. (b) Pose dependent calibration might be successfully approximated by using the mean value.

Changes in Absolute Orientation A change in orientation results from the object having a depth (Figure 4 (a)). It can be shown geometrically that the orientation error is maximal when the vector $\overline{P_1P_2}$ is roughly orthogonal to $\bar{D}$ and is purely in the vertical plane. The amount (θ) and the axis ($\hat{\kappa}$) of rotation, through a series of computations can be shown to be as in Equation (8). The bound on the rotation error is dependent only on origin mis-calibration and not on that in focal length . More importantly it is independent of the height/depth of the object (as far as it is non-planar) and its distance from the image plane. Thus F_A, F_B in Figure 8 will observe the same absolute rotation, in effect not experiencing any relative rotation. Experimental results corroborating this claim are available in Section 3.

$$|\theta| \leq \arcsin\left[\frac{\sqrt{D_x^2+D_y^2}}{f_{1z}}\right] \sim \frac{\sqrt{D_x^2+D_y^2}}{f_{1z}} \; ; \quad \hat{\kappa} = \frac{1}{\sqrt{D_x^2+D_y^2}}(D_y, -D_x, \; 0) \tag{8}$$

Error in Reconstruction of Point Features In many applications (particularly in brachytherapy), C-arms are used to reconstruct 3D point objects. This is done by obtaining multiple images at varying orientations and then using triangulation to obtain the desired intersection. In ideal circumstances, all the lines would intersect at a unique point. In practice however, calibration (and other) errors lead to non-intersecting lines. We will attempt to bound the error in this symbolic reconstruction of the point. Let point P be imaged from N different poses and reconstructed in a tracker frame F_A, which is stationary wrt P. Let the i^{th} pose have a focal spot error (in frame F_A) of $\bar{D}_i$. Without errors, each reconstructed line (l_i) would pass through P. It can be shown that due to the calibration error $\bar{D}_i$, the new line passes through a new point $\bar{P}'_A$ and undergoes a rotation ϕ.

$$\bar{P}'_A \sim \bar{P}_A + [0 \;\; 0 \;\; \frac{(\bar{P}_A \cdot \bar{D}_i)}{f_{iz}}]' \quad ; \qquad \phi \sim \frac{(\hat{l_i} \cdot \bar{D}_i) \; sin\theta_i}{f_{iz}} \tag{9}$$

where θ_i is the angle that l_i makes with the z-axis of F_A. The rotation is fairly small and can be ignored. Thus P_A is at a distance of $(\bar{P}_A \cdot \bar{D}_i)sin\theta_i/f_{iz}$ from l_i. If Q is the symbolic intersection of all l_i's, then it can be shown that Q is no further away than $(\frac{d_{max}}{f_z} sin\theta_{max})\|P_A\|$ away from any of the lines. Moreover, the reconstruction error (RE) can also be shown to be bounded by

$$RE = \|(\bar{Q} - \bar{P}_A)\| < \frac{\sqrt{2} \; d_{max}}{f_z} \; \|\bar{P}_A\| \tag{10}$$

where d_{max} is the maximum amount of mis-calibration and f_z is the minimum focal length. Thus a 10 mm focal length error causes an error less than 0.5 mm for a point at a distance of 35 mm. Note that this is the worst case error analysis and in practice the dot product in Equation (9) mutually cancels positive and negative errors, leading to extremely low reconstruction errors (Section 3).

The Optimal Choice for Calibration Parameters Since the focal spot is pose dependant, and the previous results suggest robustness to mis-calibration, choosing a constant calibration for quantitative reconstruction might be viable. In the scenario that the focal spot might vary as much as 10 mm from one pose to another, "*what constant calibration should be chosen to minimize error*"?

Let us assume that we are imaging a point P from N different poses (Figure 2 (c)). Wrt frame F_I, let the i^{th} pose have the focal spot at $\bar{f}_i = (f_{ix}, f_{iy}, f_{iz})$ and the point be at location $P_i = (X_i, Y_i, Z_i)$. Note that we assume: (a) variations in each of $f_{ix}, f_{iy}, f_{iz}, X_i, Y_i, Z_i$ & pose are independent; (b) P_i's are close to the iso-center, *i.e.* variations in X_i, Y_i, Z_i are not high. We choose a constant value of $\bar{F} = (F_x, F_y, F_z)$ for the focal spot, which will displace the point P_i to $Q_i = \mathcal{T}(\bar{f}_i, F) \cdot P_i$. The aim is to choose an $\bar{F}$ which minimizes the net variation of $\Delta Q_i = Q_i - \mu_Q$. Through a series of computations, it can be shown that

$$\mu_Q = \mu_P + \frac{\mu_z}{\mu_{fz}}(\bar{F} - \bar{\mu}_f) \tag{11}$$

$$\Delta Q_i = (P_i - \mu_P) + [\frac{\Delta Z_i}{\mu_{fz}} - \frac{\mu_z \Delta f_{iz}}{\mu_{fz}^2}]\bar{F} + \frac{\mu_z \Delta f_{iz}}{\mu_{fz}^2}\bar{\mu}_f - \frac{\Delta Z_i}{\mu_{fz}}\bar{\mu}_f - \frac{\mu_z}{\mu_{fz}}\Delta\bar{f}_i \tag{12}$$

where $\mu_Q, \mu_P, \mu_z, \mu_{fz}, \bar{\mu}_f$ are the mean values of $Q_j, P_j, Z_j, f_{jz}, \bar{f}_j$; $Z_j = \mu_z + \Delta Z_j$ and likewise for $f_{jz}, \bar{f}_j$, where $j = 1 \ldots N$. In the above calculations, the second order terms either summed to 0 due to the independence of the variables or were too small in comparison. Our choice of $\bar{F}$ should be the one that minimizes the $variance(\Delta Q) = var(\Delta Q_x) + var(\Delta Q_y) + var(\Delta Q_z)$. It should be noted that F_z scales the whole space, *i.e.* a lower value will decrease the variance, implying that the choice of $F_z = 0$ forces $var(Q_z) = 0$ by forcing all $Q_i's$ to lie on a plane. Thus $var(Q_z)$ does not provide sufficient constraints for F_z. We will first obtain F_x, F_y by minimizing the variance along x, y-axes (since there is no scaling in these directions), and then will compute F_z. Notice that the first term in Equation (12) is due to the relative movement in P, while the rest is due to an error in the calibration. Since we are interested only in the variance due to mis-calibration, we will ignore the variations in P. Minimizing $var(\Delta Q)$ and enforcing independence of f_{ix}, f_{iy} & f_{iz} gives

$$\bar{F} \;=\; \bar{\mu}_f - \frac{\Sigma_1^N \Delta f_{iz} \Delta \bar{f}_j}{\Sigma_1^N \Delta f_{iz}^2}\, \mu_{fz} \;=\; [\, \mu_{fx}, \;\; \mu_{fy}, \;\; 0\,]^T \tag{13}$$

As expected, $F_z = 0$ from above. To compute F_z, we need to impose length preserving constraints. Thus if we measure a line segment of length l in each image, use Equation (6) to derive the net length error, the minimization implies

$$F_z \;=\; \mu_{fz}(1 - \frac{\Sigma_1^N \Delta f_{iz}^2}{N \mu_{fz}^2}) \;\sim\; \mu_{fz} \tag{14}$$

Thus $\bar{F} = \bar{\mu}_f$ (the mean), which is fairly intuitive and probably in common practice. Likewise, this particular choice of F_x, F_y is also a length preserving constraint, *i.e.* it minimizes the error in lengths of line segments. Calibration error in ΔQ_i now reduces to $-\frac{\mu_z}{\mu_{fz}}\Delta\bar{f}_i$, which has a stable mean and low variance. Equation (15) gives a bound on the error when the assumed value of $\bar{F}$ is away from the mean $\bar{\mu}_f$ by a distance d. A 10 mm variation in the focal length ($var \sim 3$ mm), $P's$ roughly at the iso-center having a depth variation of 100 mm and the assumed calibration unusually away from $\bar{\mu}_f$ by 50 mm still bounds the maximum error by 0.75 mm.

Thus large and constant mis-calibration in many applications, might still provide sub-millimetric 3D quantitative measurements.

$$error \quad \leq \quad \frac{\sqrt{d^2 var(Z) + \mu_z^2 var(\|\tilde{f}\|)}}{\mu_{fz}} \tag{15}$$

2.2. Mis-calibration and its effect on Virtual Fluoroscopy

In traditional C-arm fluoroscopy, surgeons often use C-arms whenever they want to know the location of a tool inside the patient. In virtual fluoroscopy (VF), X-ray images are created synthetically with the position of the tool (tracked in 3D) superimposed on the 2D images. This not only adds to the targeting accuracy, but also reduces the radiation dose both to the surgeon and the patient. Thus VF allows the tool's position to be known without taking a new image each time.[8]

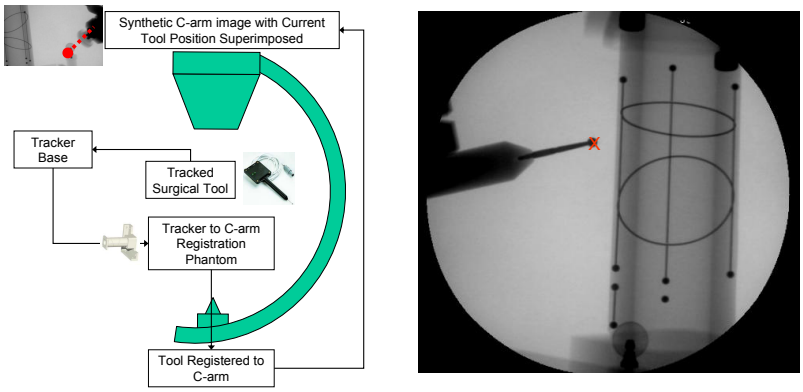


Figure 5. (a) The overall dataflow for a virtual fluoroscopy system. The current 3D tool location is registered to a previously taken C-arm image of the patient, and then rendered artificially on the image. Two such images can provide real-time 3D guidance, without the need to take more images. (b) An example screen-shot of our implemented virtual fluoroscopy system. The 'x' denotes the system's projected tool-tip position, overlaid on top of the corresponding dewarped image. For reference, the *true* location of the projected tool tip is also visible in the X-ray image. The FTRAC fiducial used to register the tracker to the C-arm is also visible.

As illustrated in Figure 5, the tool is tracked realtime in 3D using an auxiliary tracking system. The tracker is registered to the C-arm using some sort of a prefabricated precision phantom, visible both to the tracker and in the X-ray image. This allows for the projection of the current 3D tool location on a previously taken X-ray image of the anatomy, essentially creating a 'virtual' X-ray image. A prerequisite for projecting the current tool location on the image is to know the location of the X-ray image wrt the X-ray source accurately (otherwise known as C-arm calibration).

Traditionally, the C-arm is extensively calibrated with each change in pose before taking an image, which is a cumbersome procedure. A typical such phantom is illustrated in Figure 1. This naturally leads us to ask *'is it actually necessary to calibrate the C-arm for every change in pose?'*. More importantly, how precise does the calibration need to be for a *desired* accuracy in a surgical procedure. In what follows, using the mathematical framework developed so far, we study the effect of this mis-calibration on the accuracy of tool projection.

To develop the framework, we constrain the tool only to a point, which is reasonable since most tools are a collection of 3D points attached rigidly to the tracker on the tool (for ex. an insertion needle). We start with adding a calibration error of $\bar{D} = (D_x, D_y, D_z)$ in the source position wrt the image. The results in Equation

(4), (8) indicate that the primary errors are in the plane containing the true X-ray source, incorrect X-ray source and orthogonal to the image. We will refer to this plane as the *error-plane* and is illustrated in Figure 6. Thus this plane can be assumed to be the base plane with the mis-calibration $\bar{D} = (\sqrt{D_x^2 + D_y^2}, 0, D_z)$. For simplicity of calculation, we will drop the y-coefficient and assume $\bar{D} = (\sqrt{D_x^2 + D_y^2}, D_z)$.

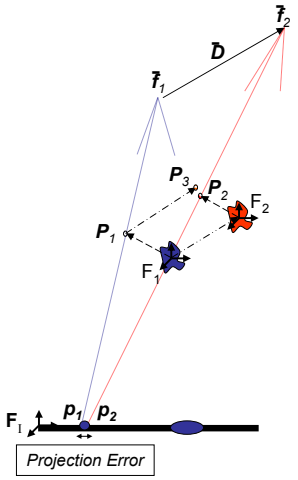


Figure 6. With the true calibration $\bar{f}_1$, the 3D point P_1 wrt F_1 projects to p_1 in the image. Due to the incorrect estimate $\bar{f}_2$, F_1 is mis-estimated at F_2, moving the point from P_1 to P_2, eventually projecting at p_2 in the image. P_3 is a *convenient* 3D point, which when projected from $\bar{f}_2$ projects on p_1. $\|p_1 - p_2\|$ represents the error in the system due to mis-calibration. Note that the error is assumed in the X-Z plane without any loss in generality.

Let F_1 be the true frame coordinate of the Tracker to C-arm registration phantom, and F_2 be the frame of the phantom as computed using an incorrect calibration. Though illustrated such, F_1, F_2 are not necessarily in the error plane. The relationship between F_2 and F_1 is as evaluated in Equation (4), (8), involving a translation in the direction of $\bar{D}$ and a rotation along the Y-axis. Let $P_1 = (a, b)$ be the true location of the tool tip wrt the phantom in frame F_I. Note that this value (when expressed in frame F_1) is independent of C-arm calibration and is computed using the tool to tracker-base transformation and tracker-base to phantom registration. Let P_2 (expressed in frame F_I) be the new location of the tool tip wrt the phantom when the frame of the phantom has been reconstructed using a mis-calibrated C-arm (F_2). Let p_1 be the true projection of the tool tip from P_1, and p_2 the mis-calibrated projection of the tip from P_2. Note that p_1, p_2 will not be the same in general, the difference representing the projection error due to mis-calibration in a VF system. We will attempt to compute an upper bound on this error.

To compute $\|p_1 - p_2\|$, we introduce another 3D point P_3, such the P_3 when projected from the mis-calibrated X-ray source, projects on the 2D point p_1. Thus points P_2, P_3 when projected using the mis-calibrated parameters, can be used to estimate $\|p_1 - p_2\|$. One example point for P_3, computed using Equation (4), is as satisfied by the equation below, where Z_1 is the Z-coordinate of point P_1 wrt F_I.

$$P_3 - P_1 = (\frac{Z_1}{f})\bar{D} \tag{16}$$

Note that P_2 wrt frame F_2 is same as P_1 wrt frame F_1. Let this vector be $\bar{v}$. If θ as computed by Equation (8) be the angle of rotation (along the new Y-axis), then $\bar{v}$ is rotated by θ along the Y-axis of the C-arm. Thus P_2 can be computed from P_1 by applying the translation and rotation of from F_1 to F_2 as given by Equations(4), (8). If Z_F is the Z-coordinate of point F_1 wrt F_I, then

$$
\begin{aligned}
P_2 - P_1 &= Translation(\bar{v}) \;+\; Rotation(\bar{v}) \\
&= [(\frac{Z_F}{f})\bar{D}] \;+\; [(b,-a)sin\theta - (a,b)(1-cos\theta)] \\
&\approx \frac{Z_F}{f}\bar{D} \;+\; (b,-a)\frac{\sqrt{D_x^2+D_y^2}}{f} \;-\; (a,b)\frac{D_x^2+D_y^2}{f^2} \\
&\approx \frac{Z_F}{f}\bar{D} \;+\; (b,-a)\frac{\sqrt{D_x^2+D_y^2}}{f}
\end{aligned}
\tag{17}
$$

Note that we have not written down the components of P_1, P_2, P_3 along the Y-axis. This is because, the affine transformation model does not make any transformations (rotation or translation) along this axis, resulting in *no change* from P_1 to P_2, P_3. From Equations (16) & (17), we get

$$
\begin{aligned}
P_2 - P_3 &= \frac{-b\bar{D} \;+\; (b,-a)\sqrt{D_x^2+D_y^2}}{f} \\
&= \frac{-b(\sqrt{D_x^2+D_y^2}, D_z) \;+\; (b,-a)\sqrt{D_x^2+D_y^2}}{f} \\
&= (0, -\frac{a\sqrt{D_x^2+D_y^2} \;+\; bD_z}{f}) \\
&= (0,\beta)
\end{aligned}
\tag{18}
$$

Note that the x-coordinate of the vector $\bar{P_2} - \bar{P_3}$ is zero. We will use this special structure to evaluate $\|p_2 - p_1\|$. For this computation, we can assume without loss of generality that F_I is located at the origin of the mis-calibrated image plane. If wrt F_I, $P_3 = (x,y,z)$ (y-coordinate included) then $p_1 = (f'/f'-z)(x,y)$, where f' is the mis-calibrated focal length. Since $P_2 = P_3 + (0,0,\beta)$, it can be shown that

$$
\begin{aligned}
\|\bar{p_1} - \bar{p_2}\| &= \|\frac{f'}{f'-z}(x,y) \;-\; \frac{f'}{f'-z-\beta}(x,y)\| \\
&= \|\frac{\beta}{f'-z-\beta}\bar{p_1}\| \\
&= (\frac{f'}{f'-z})\;(|\beta|)\;(\frac{\sqrt{x^2+y^2}}{f'-z-\beta}) \\
&\approx m|\beta|tan\phi
\end{aligned}
\tag{19}
$$

where m is the magnification factor of the 3D point and ϕ is the angle that the 3D point makes with the projection axis of the C-arm (the line joining the X-ray source and the origin, usually very close to the center of the image). Thus the error due to mis-calibration in a virtual fluoroscopy system is

$$
\begin{aligned}
\|\bar{p_1} - \bar{p_2}\| &\approx mtan\phi\frac{a\sqrt{D_x^2+D_y^2} \;+\; bD_z}{f} \\
\implies 3DError &\approx \frac{(a,b)\cdot(\sqrt{D_x^2+D_y^2}, D_z)}{f}tan\phi
\end{aligned}
\tag{20}
$$

From Equation (20) it can be noticed that the error in a VF system can be significantly reduced by: (a) bringing the tracked tool closer to the center of the image; and/or (b) using a registration phantom that can work near the tool workspace. To get a sense of the numbers, if the tool is somewhere in the middle of the C-arm workspace (magnification = 2) & projects at the corner of a 9" image intensifier ($tan\phi \approx 0.12$), then a mis-calibration as high as even 50 mm will not induce an error greater than 1 mm in a VF system, if the registration phantom is used inside a 100 mm region of the tool workspace. Alternately, even if the registration phantom needs to be kept far away (at a distance of, say, 250 mm), then keeping the projection of the tool in a 40 mm diameter region near the image center, will still produce an error less than 1 mm with a 100 mm mis-calibration. Note that a 1 mm error in the image at a magnification of 2, indicates only a 0.5 mm 3D error (when multiple images are used, as is customary). Also note that in most cases, typical mis-calibrations are predominantly in the imaging direction, in which case, it suffices if the registration phantom can be used at the same 'depth' as the tool workspace. In such cases, the dot product in Equation (20) cancels out, ensuring that no error is added in the tool guidance of the VF system.

2.3. Distortion Correction

C-arm images suffer from image distortion (as high as 5 mm), owing to the curved nature of detector (radial distortion) and the earth's magnetic field[3] (s-shaped distortion). In order to achieve quantitative surgical navigation, the distortion needs to be corrected for. This is cumbersome, more so, since the warp is pose dependent. Traditional distortion correction techniques use a calibration phantom[1–4] and characterize the distortion by fitting a high-order polynomial between the distorted image and the expected image. Though easy to use, they are cumbersome for intra-operative correction due to: 1) the interference between the patient anatomy and the phantom; and 2) increased radiation dose & surgical time when two images are used, one with and the other without the phantom. Alternately, the distortion parameters at a particular pose could be interpolated from the multiple neighboring poses taken pre-operatively,[3] but require accurate c-arm tracking. The intra-operative use of the same c-arm pose, as was used pre-operatively for distortion correction, has also been proposed,[2] though it requires the c-arm pose to be repeatable and hence could become time-consuming when collecting many images. All in all, online distortion correction has been a cumbersome problem, without a reliable/robust solution.

In line with the general theme of the paper, we observe that image distortion does not radically change between two *neighboring* poses, *i.e.* small movements if the C-arm result in small variations in image distortion. This suggests that it might suffice to use a *constant* distortion correction in a *small* region of interest (ROI). Unfortunately, the definition of a *small* ROI or *neighboring* poses is not clear and tools that can crisply characterize these ideas have not yet been developed. In what follows, we propose some simple yet powerful tools to further reinforce these ideas from a quantitative perspective.

Statistical Characterization of C-arm Distortion: C-arm distortion has been traditionally modelled using a high-order polynomial. This polynomial is computed by using a phantom grid with radio-opaque beads in a known pattern, where the observed distorted image is compared to the expected image. Among the many basis polynomials, the authors have found the Bernstein basis of order 4/5 as sufficiently robust & fast, as also observed in the literature.[2] Once the polynomials are computed, they can be used to estimate the amount of sub-pixel distortion at any given pixel in the image. This computation can be done independently for multiple pre-operative images. To study precisely the extent of statistical variations in distortion among images of a certain ROI, a principal component analysis (PCA) can be conducted on all the distortion maps. To do so, the distortion map for each image, discretised at each pixel, is organized as a single vector in a very high dimensional space. Thus each distortion map is a single point in this high-dimensional space. To quantitatively study the amounts of variation, it shall suffice to study the statistics of this collection of points. The PCA analysis can be of tremendous aid here, and can provide a linear basis for the subspace of these points, as given below.

$$\bar{D} = \bar{M} + \sum_{i=1}^{n} \lambda_i \bar{D}_i \tag{21}$$

where $\bar{D}$ is the distortion at any given C-arm pose, M is the mean distortion map of the whole space and D_i is the i^{th} strongest (principal) component (mode). Furthermore, the eigen-values (computed during PCA) associated with each principal mode indicate the variance along each associated direction. This allows for a statistical analysis of the pose-dependant C-arm distortion. It has been observed experimentally that more than 98% of the distortion can be modelled using just the first three principal modes, indicating that though complex, pose dependant C-arm distortion lies predominantly inside a 3-dimensional linear space.[9]

This experimentally determined numerical observation (for a particular C-arm) that, a small movement in the C-arm pose will result in only a small variation in image distortion, can be used to motivate two simple algorithms for distortion correction: (a) use the mean computed from Equation (21) for correcting the distortion within a certain *acceptable* region; (2) if the mean computation is unfeasible, use a *center* image that is geographically *close* to the mean image. Intuitively, if an intra-operative C-arm image is desired *near* a certain pre-calibrated pose, then compute a mean distortion *near* that particular pose. Furthermore, from a quantitative 3D guidance perspective, the use of the pre-operative distortion map itself might suffice in some cases.

3. PHANTOM EXPERIMENTS AND RESULTS

Validity of the Model: Equations (4) & (8) give the translation and rotation transformations as predicted by the affine model, the accuracy of which would furnish the validity of the model. We used the FTRAC fiducial (Figure 7), a small image-based fluoroscope tracking fiducial, which (given the calibration) can track a C-arm with an accuracy of 0.5 mm in translation and 0.65^o in rotation.[10] The fiducial was imaged using a Philips Integris V3000 fluoroscope and the true calibration read off the machine display. The images were not corrected for distortion. The pose of the fiducial (wrt to F_I) was first evaluated using the correct calibration, and then with the mis-calibrated parameters. The difference between the pose change predicted by the equations and the one computed using the non-linear pose estimation software, is displayed in Figure 8 (a) as a function of maximum calibration error. A total of $40,500$ points (from total 75 poses) with randomly added mis-calibration were used. Even when mis-calibration is as high as 25 mm, the model can predict the rotation-axis with an accuracy of 4^o, amount of rotation by 0.1^o and translation under 5 mm. For extreme mis-calibrations the translation error linearly increases, while rotation is still stable. Thus the model seems to predict with an acceptable accuracy.

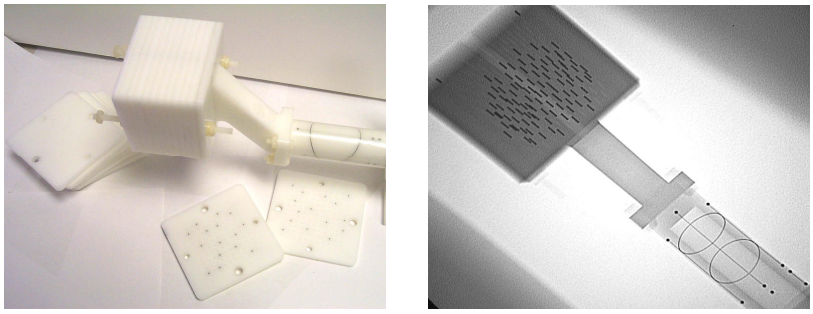


Figure 7. An image of the seed phantom attached to the FTRAC fiducial (left). The seed phantom can replicate any implant configuration, using the twelve 5 mm slabs each with over a hundred holes. A typical X-ray image of the combination (right).

Accuracy of C-arm Tracking: The FTRAC fiducial was mounted on a 0.02^o accuracy rotational turntable, while the fluoroscope was kept stationary. The turntable was rotated by known precise amounts (ground truth) and images were taken. The relative poses were also computed using the pose estimation software. The accuracy in the estimation of C-arm motion is given by the difference between the computed relative pose and the true relative pose. The tracking accuracy is plotted in Figure 8 (b) as a function of mis-calibration. Even a high mis-calibration of 150 mm adds no *additional* error in C-arm motion estimation, fixing the value at 0.45 mm in

translation and 0.6^o in rotation. An unusually high mis-calibration of 400 mm also only marginally decreases accuracy. Thus, mis-calibration does not increase the error of C-arm tracking .

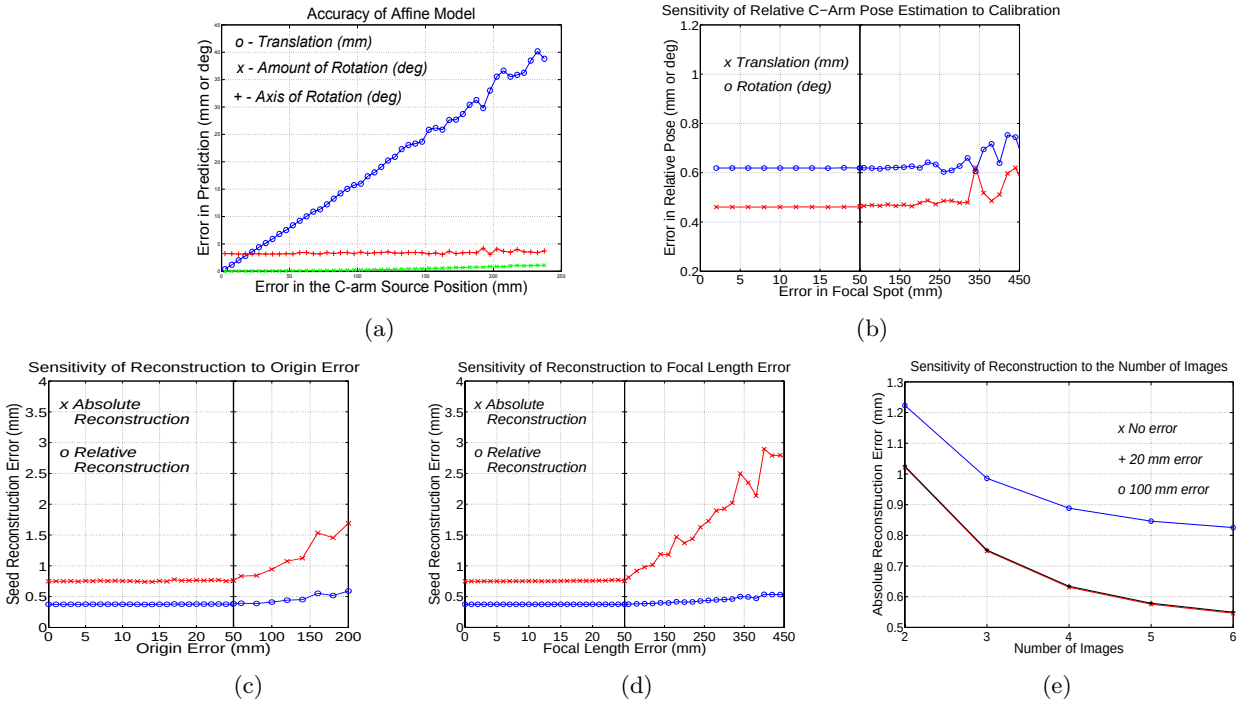


Figure 8. Note the scale variation in x-axis. (a) An affine transformation is able to predict the movement of 3D objects due to mis-calibration; (b) C-arm tracking is insensitive to mis-calibration; 3D Reconstruction is insensitive to mis-calibration in (c) origin ; (d) focal length up to 50 mm, beyond which it starts to linearly drift away from the tracking fiducial. Notice that the shape of the implant (relative err) is barely altered; (e) 3D reconstruction error decreases with an increase in images used.

3D Quantitative Reconstruction using Multiple Images: In addition to tracking a C-arm, it is equally important that multiple objects in the field of view (eg. vertebrae and screws) be reconstructed accurately relative to each other. In order to validate our hypothesis that 3D reconstruction might not be sensitive to mis-calibration, we use an accurate acetol phantom (Figure 7) having 100 dummy radioactive seeds, approximating a brachytherapy implant (Figure 7). The true 3D coordinate of each seed wrt the fiducial is known by rigid attachment. The C-arm is tracked using the FTRAC fiducial and the 3D seed coordinates are computed by triangulation (an algorithm called MARSHAL is used to establish correspondences). The difference between the computed and the true seed location gives us the 3D reconstruction error for each seed (wrt fiducial). The relative reconstruction error removes any consistent shift reflecting any change in shape. These errors are plotted as a function of mis-calibration in Figure 8 (c), (d). The reconstruction error is insensitive to mis-calibration in origin and focal length errors of up to 50 mm. The shape of the implant is stable even for large calibration errors. Figure 8 (e) shows a drop in reconstruction error as the number of images increase. Thus mis-calibration does not decrease reconstruction accuracy.

3D Virtual Fluoroscopy: We implemented an experimental setup in order to evaluate the sensitivity of a VF system to C-arm mis-calibration, the main components of the system being: (a) A GE 9600 C-arm; (b) a precise C-arm distortion correction and calibration toolkit (implemented in Matlab); (c) a Polaris tracker; (d) a calibrated Polaris pointer (calibrated using the NDI ToolViewer); (e) the FTRAC fiducial.

Distortion correction and precise C-arm calibration is performed for the C-arm images using the phantom shown in Figure 1. The phantom is a typical two plate contraption that can be attached to the intensifier. The bottom plate with the evenly spaced metal BB's corrects for any image distortion. Each straight line on the upper plate, along with its projected image-line constraints the X-ray source to a 3D plane. The least square intersection of these ten planes provides an accurate estimate of the source location wrt the image (C-arm calibration). In order to display the current location of the tool in a VF system, the tracker (Polaris) needs to be registered to the C-arm. By virtue of its design, the FTRAC fiducial can register itself to the C-arm (in every image that it is present). Thus, the Polaris to C-arm registration can be achieved by registering the FTRAC fiducial to the Polaris. To realize this, the tracked Polaris pointer was used to digitize 4 straight lines on the FTRAC fiducial, the exact location of which was precisely known from the design file. A non-linear optimization algorithm aligns these straight lines, essentially yielding the FTRAC-Polaris transformation. Note that this registration is C-arm pose independent.

At any given pose of the C-arm, the calibration phantom is used to dewarp the image and to compute the *true calibration*. The phantom is then detached and a new image is taken with both the pointer-tip and the FTRAC fiducial visible in the image (attached steadily using clamps). The 3D pointer location is measured using the Polaris. The location of the pointer tip can now be projected in the X-ray image using the transformation

$$p_{proj} \;=\; {}^{I}\mathcal{P}_{C} \;\cdot\; {}^{C}F_{F} \;\cdot\; {}^{F}F_{P} \;\cdot\; {}^{P}F_{T} \;\cdot\; p_{tip} \tag{22}$$

where p_{tip} is the location of the tool-tip wrt the tool frame (F_T), ${}^{P}F_{T}$ is the measured transformation from the tool frame to the Polaris frame (F_P), ${}^{F}F_{P}$ is the pre-computed transformation from the Polaris frame to the FTRAC frame (F_F), ${}^{C}F_{F}$ is the transformation from the FTRAC to the C-arm frame (F_C) computed for every image, ${}^{I}\mathcal{P}_{C}$ is the projective transformation that converts any 3D point to its projection in the 2D image (F_I), and p_{proj} is the *(computed)* VF projected tool-tip location in the image. Note that ${}^{I}\mathcal{P}_{C}$ and ${}^{C}F_{F}$ will change from one C-arm mis-calibration to another. If p_{seg} is the *true* projection of the tool-tip, segmented from the X-ray image, then $\|p_{proj} - p_{seg}\|$ provides the Euclidean error estimate in the VF system.

A total of 75 such X-ray images were acquired by randomly varying the C-arm pose (15 poses) and the location of the tracked tool wrt the FTRAC fiducial (5 per pose). For any known amount of artificially added mis-calibration, the series of frame transformations was used to project the pointer tip on each image (p_{proj}) and then compared to its corresponding p_{seg}. Thus to experimentally study the sensitivity of a VF system to C-arm mis-calibration, $\|p_{proj} - p_{seg}\|$ is plotted as a function of mis-calibration in various graphs as shown in Figure 9.

It can be seen that the overall average accuracy of out VF system without any mis-calibration is about 2.05 mm (STD 1 mm) in the 2D X-ray images. This indicates a 3D targeting error of about 1 mm at our magnification (0.4), which is reasonable owing to the accuracies of (a) the polaris (0.4 mm); (b) FTRAC (0.6 mm); & (c) pointer calibration (0.2 mm).

Figure 9 (a) - (c) plot the mean/std of the VF system as a function of uniformly distributed mis-calibration. Each graph is computed from a total of 40,500 random points (75 images, 12 intervals, 45 random runs per image per interval). Note that the three plots remain extremely stable & indicate a similar pattern. The mis-calibration in focal length is slightly more critical than the ones in the origin. Furthermore, the graphs indicate that a mis-calibration, as high as even 100 mm, does not result in any significant loss in accuracy. To decouple the mis-calibration errors with any intrinsic error in our VF system, in Figure 9 (d) we plot the distance of the mis-calibrated projection to the projection obtained by using perfect calibration. Thus offsetting for the inherent errors in the system, it helps us understand the *variation* in the tool projection, purely as a function of C-arm mis-calibration. It can be observed that on an average, the tool tip projection moves by less than a pixel (0.44 mm) for mis-calibration as high as even 100 mm. Figure 9 (e) plots the error in the VF system as a function of the distance between the tool tip and the FTRAC fiducial used for the registration. As indicated by Equation 20, there is a slight increase in the error of the mis-calibrated VF system as the distance between

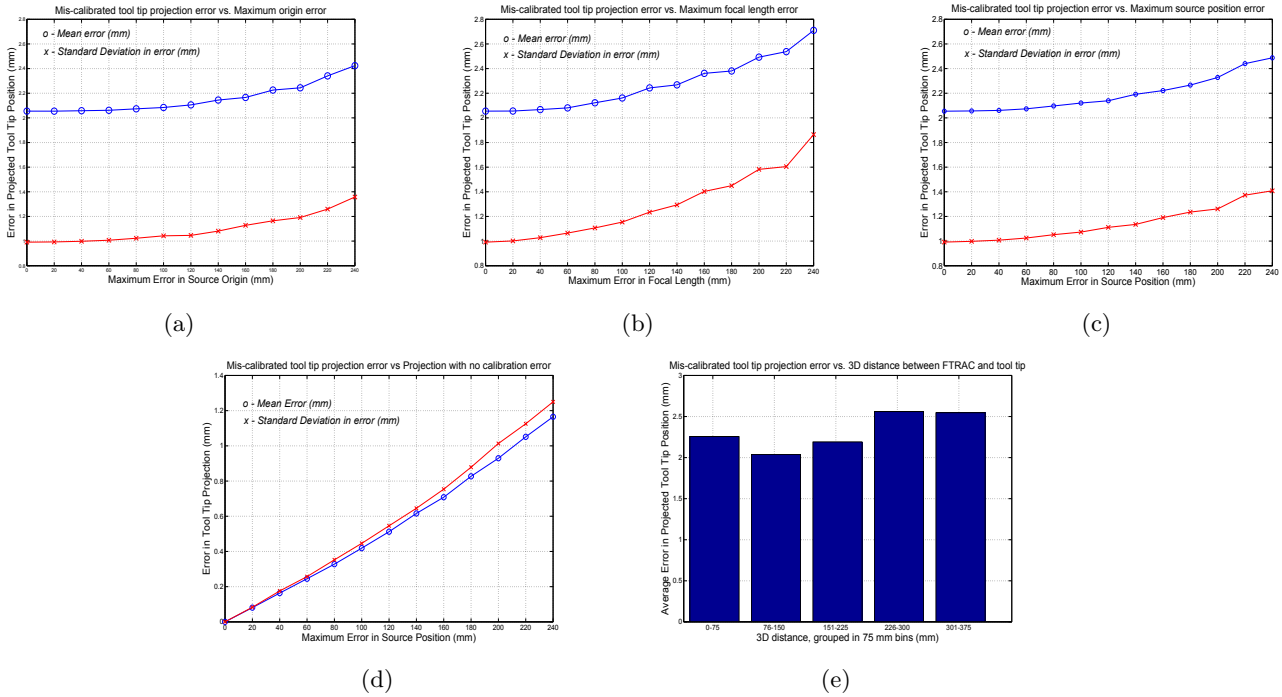
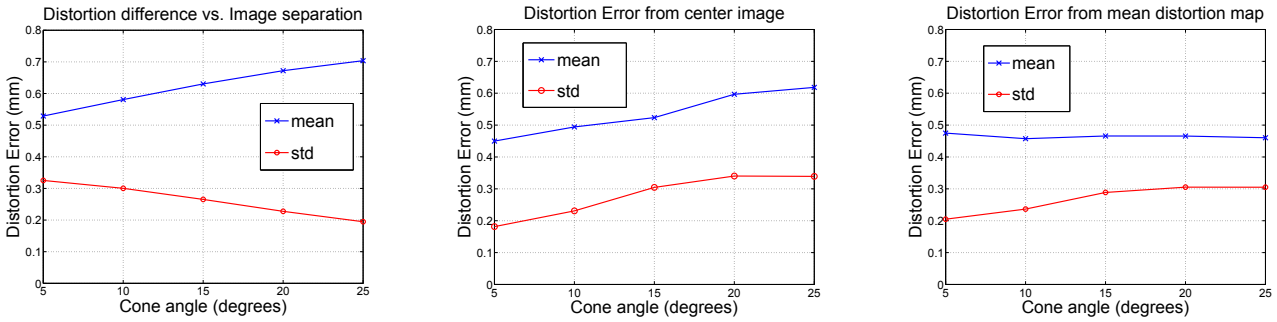


Figure 9. The error in a virtual fluoroscopy system using a total of 40,500 samples plotted as a function of mis-calibration in (a) C-arm origin ; (b) C-arm focal length ; (c) origin & focal length together. It can be noticed that VF guidance is practically independent of mis-calibration. (d) The *additional* error in the VF system is plotted as a comparison to the projection with the perfect calibration at each image. It can be noticed that the *additional* error due to mis-calibration is not significant, indicating that the primary sources of error are from other sources like tool tracking or tracker-fluoroscope registration. (e) The error in the VF system plotted as a function of the distance between the anatomy (tool-tip) from the registration fiducial. Typical patient thickness of 300 mm does not add any significant error to the system.

the tool-tip and the fiducial goes over 300 mm. Nevertheless, this increase is only marginal (less than a pixel), indicating a significant amount of robustness to the tool-tip position. Both graphs in Figure 9 (d)-(e) use the same data as used in (c). Thus, it can be concluded that mis-calibration as high as 50 mm does not add any significant amount of additional error in a VF system. Moreover, this additional error can tolerate a 350 mm distance from the tracked-tool to the registration fiducial, a number representing typical patient sizes.

Distortion Correction: With the distortion correction phantom attached to the intensifier, 86 C-arm images in a 25^o cone around the AP-axis were acquired at a 5^o separation. The average distortion in the images was found to be 3.31 mm with a standard deviation of 1.41 mm. Figure 10 (a) plots the average variation in the image distortion from one pose to another, as a function of angular image separation. Based on the two methods suggested in Section 2.3, Figure 10 (b), (c) plot the performance of each algorithm as a function of the repeatability of C-arm pose. The center-image distortion correction function reduces the average distortion in the image to an average value of 0.63 mm (STD 0.33 mm), using just a single center image in a 25^o cone. The mean-image distortion correction function performs better, reducing the residual distortion error to an average of 0.48 mm (STD 0.3 mm). Though the center-image correction method is less accurate than the mean-image correction method, it has the convenience of using just a single pre-operative image *near* the intended region of C-arm use. Thus the results indicate that using a single distortion correction image can reduce the average distortion error in the image from 3.31 mm to 0.63 mm when the intra-operative C-arm pose can be repeatability positioned with 25^o accuracy, which is sufficiently acceptable for many applications. The use of a mean-image by utilizing multiple images *close* to the intended region of use, can further decrease this error to 0.48 mm, though

at the expense of collecting multiple images in the region. Note that these numbers for residual distortion errors are close to the resolution of the X-ray images (0.44 mm).



(a) (b) (c)

Figure 10. (a) The variation in C-arm distortion from one pose to the other as a function of angular separation. We notice that distortion is *continuous* and does not change radically near a certain neighborhood. (b) The performance of using the geographic center-image for distortion correction of all images in a region. The residual error in distortion is shown as a function of pose repeatability (cone half-angle). Only a single image is needed for distortion correction in the whole neighborhood of interest. (c) The performance (residual error) of using a computed mean-distortion map for the correction of all images inside a region. Multiple pre-operative distortion correction images in the region will be needed to compute an estimate of the mean distortion.

4. CONCLUSION

We modelled the the effects of mis-calibration on as an affine transform, and proved its validity experimentally on a variety of common fluoroscopy procedures involving C-arm tracking, 3D reconstruction or virtual fluoroscopy. We have derived bounds on the amount of scaling, translation and rotation error. For pose dependant calibration, we proved that using the mean calibration minimizes the reconstruction variance. Phantom experiments with a radiographic fiducial indicate that C-arm tracking is insensitive to mis-calibrations. We also showed that mis-calibration up to 50 mm adds no additional error in 3D reconstruction of small objects, beyond which the reconstructed objects begin to drift wrt the fiducial, while still retaining the shape. For the case when external trackers are used in conjunction with C-arms, we showed that mis-calibrations of up to 50 mm are still acceptable if the tool-projection is kept in the center of the image or if the tracker-fluoroscope registration phantom is kept close to the patient anatomy. Experimentally, we showed that virtual fluoroscopy can tolerate mis-calibrations as high as 100 mm. Furthermore, to address the problem of pose dependant distortion correction, we propose the use of a mean-image, reducing the distortion close to 1 pixel error. In conclusion, a significant family of quantitative fluoroscopy applications involving localization of small markers can function without cumbersome on-line calibration. A constant loose calibration might suffice.

5. ACKNOWLEDGMENTS

This work has been supported by DoD PC050170 - Prostate Cancer Research Program (2005) pre-doctoral Traineeship Award, NIH SBIR 1R43CA099374 01, NSF EEC-9731478.

REFERENCES

1. R. Hofstetter, M. Slomczykowski, M. Sati, and L. Nolte, "Fluoroscopy as an imaging means for computer-assisted surgical navigation," *CAS* **4(2)**, pp. 65–76, 1999.
2. J. Yao, R. H. Taylor, and et al, "A c-arm fluoroscopy-guided progressive cut refinement strategy using a surgical robot," *Comput Aided Surg* **5(6)**, pp. 373–90, 2000.

3. R. Fahrig and et al, "Three-dimensional computed tomographic reconstruction using a c-arm mounted XRII: correction of image distortion," *Med Phys.* **24(7)**(-), pp. 1097–106.
4. H. Livyatan, Z. Yaniv, and L. Joskowicz, "Robust automatic c-arm calibration for fluoroscopy-based navigation: A practical approach," in *MICCAI*, pp. 60–68, 2002.
5. J. Siewerdsen, M. Daly, G. Bahar, D. Moseley, G. Bootsma, S. Chhabra, D. Jaffray, and J. Irish, "Multimode c-arm fluoroscopy, tomosynthesis, and cone-beam ct for image-guided interventions: from proof of principle to patient protocols," in *SPIE Medical Imaging; Physics of Medical Imaging*, Feb 2007.
6. R. Kumar and A. Hanson, "Sensitivity of the pose refinement problem to accurate estimation of camera parameters," in *ICCV90*, pp. 365–369, 1990.
7. E. Malis, "Visual servoing invariant to changes in camera intrinsic parameters," *IEEE Transaction on Robotics and Automation* **20**, pp. 72–81, Feb 2004.
8. J. C. Eyke and et al., "Computer-assisted virtual fluoroscopy," *The University of Pennsylvania Orthopaedic Journal* **15**, pp. 53–59, 2002.
9. G. Chintalapani, A. Jain, and R. Taylor, "Statistical characterization of c-arm distortion with application to intra-operative distortion correction," in *SPIE Medical Imaging; Visualization, Image-Guided Procedures, and Display*, Feb 2007.
10. A. Jain, T. Mustufa, Y. Zhou, E. C. Burdette, G. Chirikjian, and G. Fichtinger, "A robust fluoroscope tracking (ftrac) fiducial," *Med Phys* **32**, pp. 3185–98, Oct 2005.