A new technique for multi-modal 3D image registration

Gjenna Stippel[1], Jim Ellsmere[2], Simon K. Warfield[2], William (Sandy) M. Wells[2], and Wilfried Philips[1]

[1] Ghent University, Department TELIN, Gent, Belgium
[2] Harvard Medical School, Brigham & Women's Hospital, Boston, MA USA
http://www.gjenna.com

Abstract. In this paper we address the problem of multi-modal co-registration of medical 3D images. Several techniques for the rigid registration of multi-modal images have been developed; in one of those the Kullback-Leibler distance is used to align 2D-3D angiographic images [1]. In this paper we investigate the performance of this technique on the registration of pairs of 3D CT/US images. We study the effects of various image perturbations on the performance of the registration, and obtain promising results.

1 Introduction

Registering intraoperative ultrasound (US) to preoperative images like computed tomography (CT) or magnetic resonance (MR) remains a challenging problem in the field of image guided surgery. For many surgical applications, ultrasound provides adequate clinical information to carry out the required intervention. There are some applications, however, where being able to interpret the ultrasound images in the context of the higher quality preoperative imaging has been shown to be helpful [2].

In this paper, we introduce a new rigid registration technique that aligns US and CT images by applying linear transformations to the coordinate system in order to minimize the Kullback-Leibler distance (KLD) between the observed joint intensity distribution, and a reference distribution representing the prior knowledge of the expected joint intensity histogram when the images are aligned properly [1].

We calibrate our KLD registration by taking an ultrasound scan, and manually aligning it with a CT scan to establish the fixed reference distribution. Afterwards, as more 3D ultrasound acquisitions are acquired over the course of the surgery, this calibration does not need to be repeated. Hence, the cost of the calibration is amortized over all the follow up scans during the procedure. If there are many scans, the calibration cost is a small percentage of the total effort.

The performance of the method and its robustness under various image perturbations is studied.

2 The registration technique

The proposed technique is an intensity based registration algorithm. As said, in a first calibration step we align one image pair consisting of an image of the same organ from each modality manually. From this initial manual alignment a joint intensity histogram $P(g_1, g_2)$ with $0 \leq g_1, g_2 \leq 255$ is constructed as follows:

Let I_1 and I_2 be the intensity values of the two images, and let X_1 and X_2 be their image domains respectively. Since the images are aligned, we have $X_1 = X_2$. Let $J := \{(x, i_1(x), i_2(x)) | x \in X_1\}$, $(i_1 \in I_1,\ i_2 \in I_2)$. Now the joint intensity histogram P is the (two-dimensional) histogram of J, i.e., $P(g_1, g_2) := |\{(a, b, c) \in J | b = g_1 \wedge c = g_2\}|$ $(0 \leq g_1, g_2 \leq 255)$, where $|S|$ denotes the cardinality of the set S.

This histogram serves as a reference distribution in all further alignments. To register the new image automatically, we can iteratively translate and rotate one of the two images (the "moving image") while keeping the other image fixed (the "fixed image"), thereby aiming to determine the rigid transformation for which the resulting joint intensity distribution corresponds most with our reference distribution. The measure we use to quantify the similarity of the two histograms is the so-called Kullback-Leibler distance.

2.1 Kullback-Leibler distance

Let the reference distribution be denoted by P_{ref} and the observed joint intensity distribution after the transformation T has been applied to the moving image by P_T. Now the Kullback-Leibler distance between the two distributions is defined as follows:

$$D(P_T || P_{\text{ref}}) := \sum_{i_1, i_2 = 0}^{255} P_T(i_1, i_2) \log \frac{P_T(i_1, i_2)}{P_{\text{ref}}(i_1, i_2)}.$$

The Kullback-Leibler distance is not a distance in the mathematical sense, since in general it is not symmetric: $D(P_{\text{ref}} || P_T) \neq D(P_T || P_{\text{ref}})$ and thus does not satisfy one of the three criteria any distance has to satisfy. Two other distance properties are met and are the most important for our purpose [3, 4]:

1. $D(P_T || P_{\text{ref}}) \geq 0$,
2. $D(P_T || P_{\text{ref}}) = 0$ if and only if $P_T = P_{\text{ref}}$

These two properties make the Kullback-Leibler distance a useful measure to determine how far a particular alignment of the two images is from our "reference alignment" obtained from the manual calibration. For an image pair that is identical to the image pair used in the manual calibration, we observe that when these two images are perfectly aligned, the joint intensity histogram coincides exactly with the reference histogram, so in that case the KLD equals zero; otherwise the KLD is strictly positive.

The idea behind the registration technique is thus, to find a transformation T_0, acting on the moving image, that minimizes the KLD between the joint intensity distribution P_{T_0} and the reference distribution P_{ref}. Or, in formula:

$$T_0 = \arg\min_T D(P_T || P_{\text{ref}}).$$

In this paper we restrict ourselves to translations only because of their computational simplicity. The way to deal with the problem when we include rotations also is exactly the same; we only need to conduct our search in a six-dimensional space then, instead of a three-dimensional one.

Moreover, registration of volumetric images using translations exclusively has practical applications: if the images are taken under approximately the same angle with respect to the skin surface and the imaged organs do not rotate with respect to the skin surface themselves either (e.g., as with the brain), rotations have no significant impact on the registration. So in this case it is sufficient to use translations only.

The strategy we follow in the technique to find the optimal alignment (i.e., the one closest to the reference alignment) is as follows:

We keep the CT scan fixed during the entire procedure. The US image can be translated in any direction. Using a steepest descent optimization method we search, by iterative translations, the position for the moving image at which the KLD between the corresponding joint intensity distribution and the reference distribution is minimal.

Several questions arise. How robust is the method against "image variability" with respect to:

- geometric deformations; the brain is fairly rigid but it still differs in shape somewhat from one scan to another.
- overall brightness; due to different machine settings or anatomical differences, like the amount of hair an infants has on his head, the US images can differ in overall brightness from one scan to another.
- noise; even the slightest difference of the head of the transducer influences the speckle noise in the US image.

All these questions will be addressed in Section 6.

KLD has been previously used by Chung et al. [1] to do 2D-3D registration of angiographic image data.

3 Dataset

To generate our test set, we used a 3D abdominal phantom (CIRS, Norwalk, VA). The CT scan (512x512, 1mm slices) was acquired using the Somatoform Plus4 (Siemens Medical Systems, Iselin, NJ). Two 3D US images of the liver (left lobe, right lower lobe) were generated using the Stradx software (Cambridge University, Cambridge, UK). The 3D US system consists of a Lynx ultrasound unit (BK Medical Systems, Wilmington, MA) and a miniBIRD tracking system (Ascension Technology, Burlington, VT). For the calibration the CT and US images were manually aligned using the Slicer software (Brigham & Women's Hospital, Boston, MA). In Figure 1 two 2D sections that are aligned, one of the

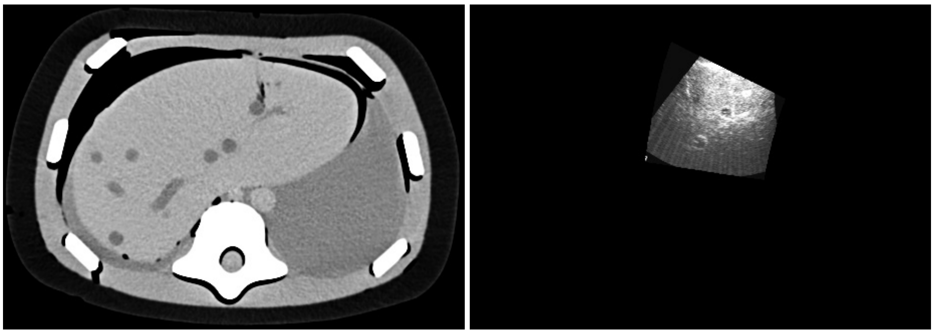


Fig. 1. 2D section of volumetric CT and US scans as used in the experiments

CT scan and one of the US scan, are shown. The size of one 3D US image is: 460x324x18 voxels.

4 Image perturbations

In order to study the effects of the aspects mentioned in Section 2.1 on the proposed technique, we have perturbed the US images in various ways:

Geometrical distortion:

- we rescaled the image to 90 % of the original size in the x-direction

Brightness:

- we applied gamma-correction (with $\gamma = 0.5$) on the image to make it darker

Noise:

- we applied a Lee filter to the image [5], using a 7×7 window, and an estimated-noise-variance parameter of 250
- we applied a median filter to the image, using a 5×1 window. Visually this choice of the window size yields the best results regarding speckle suppression without blurring the image significantly
- we added multiplicative (speckle) noise to the image artificially. If we call the image I, we added the noise using the equation $J = I + n * I$, where n is uniformly distributed random noise with mean 0 and variance 0.1 or 0.3

With all these images we performed the experiment as explained in Section 5.

5 Experiment

The experiment we performed was as follows: from the manual alignment of the US image of the left lobe of the liver and the CT scan (see Section 3) we constructed our reference joint intensity histogram. We keep the CT scan fixed during the entire procedure. The US image is moved from 100 pixels to the left to 100 pixels to the right in the x-, y- and z-direction with steps of 4 pixels. At each position a new joint intensity histogram is constructed, and the KLD between this joint intensity histogram and the reference histogram is calculated.

In Figure 2 the resulting graphs of the KLD between the unprocessed US image of the left lobe and the CT scan are displayed, when translated in the x-, y- and z-direction. In the left subfigure the translations have been performed with steps of 4 pixels, in the right subfigure with steps of 1 pixel. This demonstrates the typical shape of a KLD graph, with its minimum at the correct position.

Our ultimate goal in this work is to register more US images that are taken during the course of the surgery with the CT scan using the same reference distribution acquired from the calibration. Since, obviously, these will not be exactly the same, we repeated the experiment described above with the artificially perturbed US images (as explained in Section 4). We translate the perturbed US images with respect to the original CT scan, construct a new joint intensity histogram in every step again, and calculate the KLD between this joint intensity histogram and our standard reference distribution. In these cases we only do the translations in the x-direction from 25 pixels to the left to 25 pixels to the right, with steps of 1 pixel.

We wish to see the influence of these perturbations on the performance of the registration with respect to:

1. Accuracy
2. Speed of convergence
3. Robustness

As a measure of the accuracy we take the position of the minimum of the KLD graph. Since none of the perturbations we perform involves any translation of the image, the optimal alignment is at exactly the same position as for the original image (i.e., at 0 pixels translation). So, any number of pixels the minimum of the KLD graph corresponding to a perturbed image is next to this position, can be considered a misalignment.

The speed of convergence and the robustness are estimated by the "pointedness" of the graph. We see in Figures 3 and 4 that in general, the shape of the KLD graphs is as follows: there is one global minimum, and in the direct neighbourhood of that the graph has a "V"-shape. With the "pointedness" of the KLD graph we mean the acuteness of the angle of this "V". As a way of measuring it , we measure the width of this "V" at various heights:

If m is the x-value at which the KLD graph is minimal, then we define $b := \frac{\text{kld}(m-1)+\text{kld}(m+1)}{2}$ as the "base point". We define the base point like this, because the value of the KLD graph at the global minimum can sometimes be significantly lower than the values at the two points adjacent to it.

We then measure the width of the KLD graph at the heights: $b,\ b+0.001,\ b+0.002, \ldots,\ b+0.008$. If $t > 0$, then the width w_t at height t is:

$$w_t := \max\{x|\text{kld}(x) \leq t\} - \min\{x|\text{kld}(x) \leq t\}.$$

In Figure 4 we plotted $w_b,\ w_{b+0.001}, \ldots,\ w_{b+0.008}$ for the KLD graphs of all perturbed US images.

6 Results

We performed the experiments as explained in the text. Figure 2 consists of two subimages. In the left subimage the KLD graphs are displayed for when we move the images in the x- and the y-direction from 96 pixels to the left of the optimal alignment to 96 pixels to the right with steps of 4 pixels. In the right subimage the same KLD graph for the translations from 25 pixels to the left of the optimal alignment to 25 pixels to the right (with steps of 1 pixel) and in the z-direction from -18 to +18 pixels of the optimal alignment are displayed. These figures illustrate what (a one dimensional intersection of) the KLD graphs usually look like.

Figures 3 and 4 show us the KLD graphs of the various preprocessed images. In these cases, we only move the images in the x-direction.

The first thing we can conclude from these graphs is that the translation value (x-value) at which the KLD reaches its minimum remains 0 in all cases. (In our experiments 0 is the correct alignment). This demonstrates that with regard to finding the correct alignment the proposed technique is robust against image variability in small geometrical deformations, brightness and noise level.

In Figure 4 we compare the pointedness of the KLD graphs, as explained in Section 5. In general a more pointed KLD graph, i.e. a graph in which the angle that the minimum and the surrounding function values make is more acute, is preferable. Since in this case the KLD function values approach the global minimum more rapidly if the x-value approaches 0, this minimum is approached quicker with a steepest descent method. Furthermore, if the KLD graph is less pointed, i.e. the shape of the graph is more shallow, disturbances by noise are likelier to create local minima. Local minima are undesirable since they can attract the optimizer (e.g. steepest descent method) to the wrong optimum, thereby failing to align the images properly.

In the graphs in Figure 4 the shape of the KLD graphs reveals itself in the slope of the corresponding curves: when the KLD graph is shallow, the corresponding curve in Figure 4 is very steep, since the width of the KLD graph increases fast with the height. A more pointed KLD graph results in a less steep curve in Figure 4.

We can see from the graphs in the right subfigure in Figure 4 that a small geometrical deformation hardly has any influence on the registration performance. The corresponding graphs remains very close to that of the original image. Median filtering does not have much effect either.

Furthermore after applying the gamma correction, and making the image darker, the KLD graphs becomes more pointed, thereby indicating a faster registration. This is not unexpected: by lowering the brightness of the image a lot of speckle noise is removed while mainly the important structures remain visible.

Lee filtering the US image and adding speckle noise results in slightly more pointed KLD graphs.

7 Conclusion

In this paper we introduced a novel technique for the rigid registration of volumetric multi-modal medical images. We investigated its performance on a test set of a pair of CT and volumetric US images of an abdomen phantom. The proposed technique successfully succeeds in aligning the US-CT pair, also after the various perturbations of the US image. We are currently conducting further tests on human images to assess its suitability for clinical application.

8 Acknowledgements

This research was supported by the following grants: NSF ERC grant under Johns Hopkins Agreement 8810274, NIH grants P41RR13218, P01CA67167.

References

1. Albert C.S. Chung, William M. Wells III, Alexander Norbash, and W. Eric L. Grimson, "Multi-modal image registration by minimising kullback-leibler distance," in *Proceedings of MICCAI*, 2002, pp. 564–571.
2. J. Ellsmere, J.A. Stoll, D.W. Rattner, D. Brooks, R. Kane, W.M. Wells III, R. Kikinis, and K.G. Vosburgh, "Integrating preoperative ct data with laparosopic ultrasound images facilitates interpretation.," *Society of American Gasrointestinal Endoscopc Surgeons*, vol. 17, pp. S296, 2003.
3. T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., 1991.
4. S. Kullback, *Information Theory and Statistics*, Dover Publications, Inc., 1968.
5. J. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no. 2, pp. 165–168, 1980.

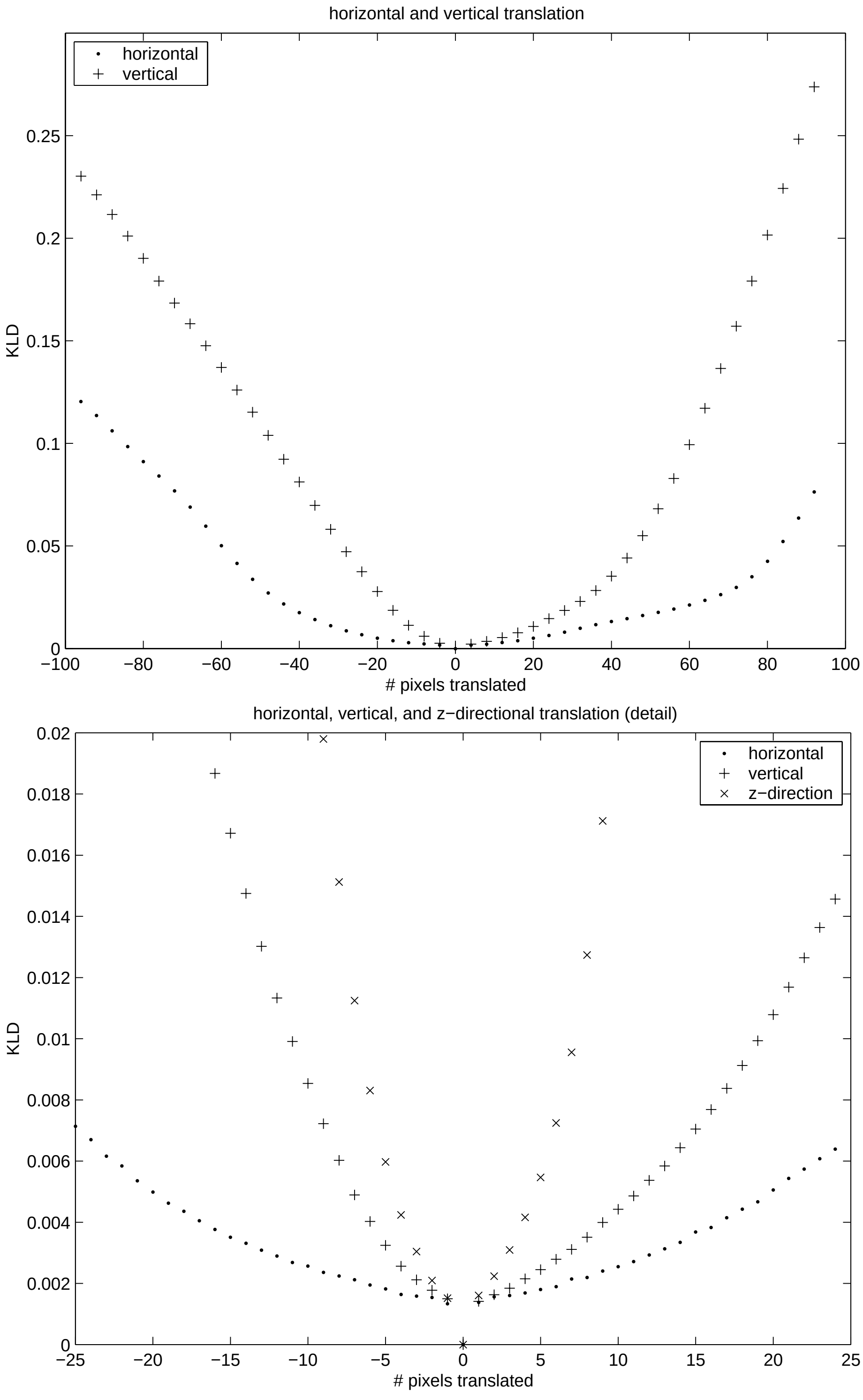


Fig. 2. KLD graphs of translations of the unprocessed US image of the left lobe

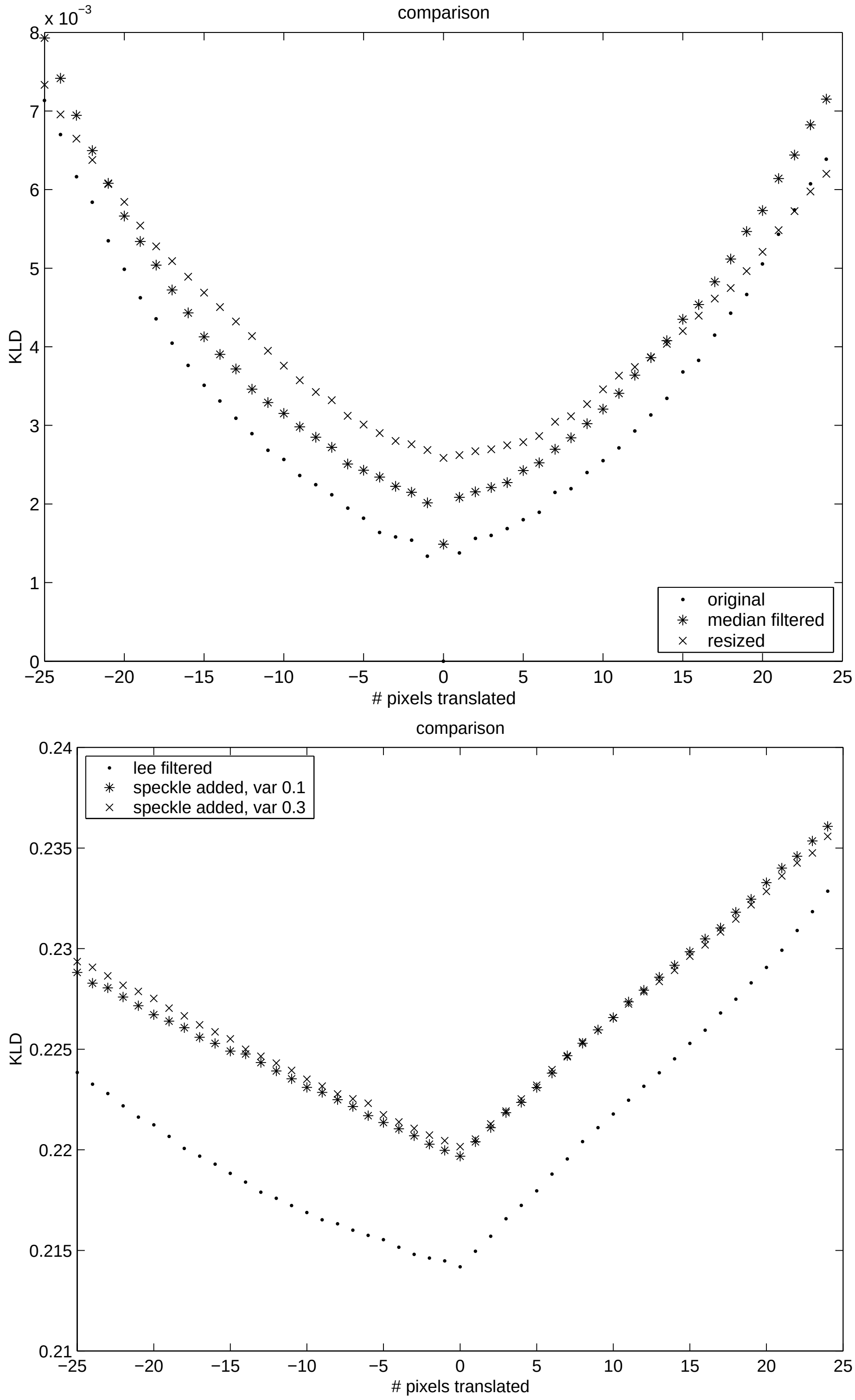
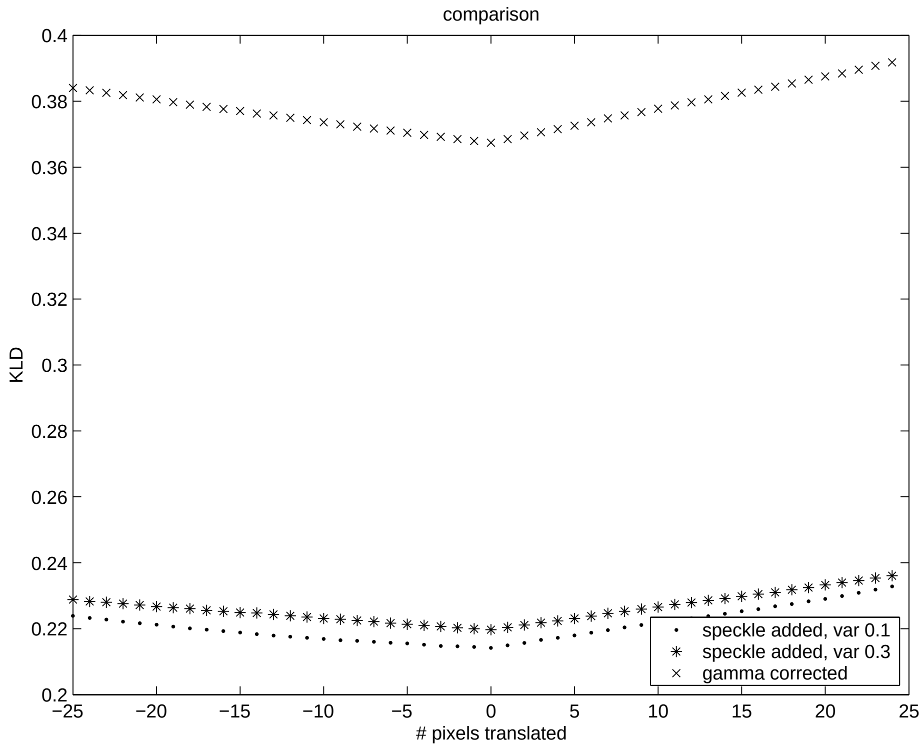


Fig. 3. KLD graphs of the perturbed images of the US image of the left lobe



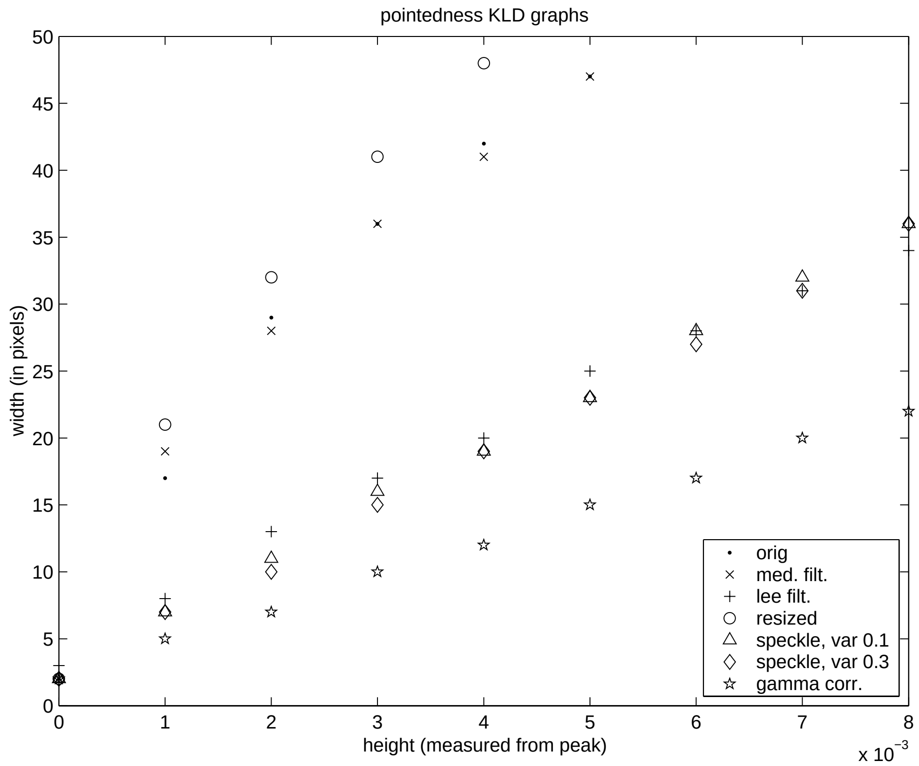


Fig. 4. Left: KLD graphs of the perturbed images of the US image of the left lobe **Right:** Pointedness of KLD graphs