FAST MUTUAL INFORMATION BASED REGISTRATION AND FUSION OF REGISTERED TOMOGRAPHIC IMAGE DATA

Peter Hastreiter, Jörg Freund, Günther Greiner, Thomas Ertl

Universität Erlangen,
Lehrstuhl für Graphische Datenverarbeitung (IMMD 9),
Am Weichselgarten 9,
91058 Erlangen, Germany
hastreiter@informatik.uni-erlangen.de

ABSTRACT

Registration of medical image data of different modalities and the consecutive fusion of registered data are the basis for visualization in medical imaging. The applied alignment procedure is based on the information theoretic approach of maximizing mutual information of two arbitrary images of different modalities. The implementation of the histogram based approach, as suggested by Collignon et al. [1], showed that very good matching results for various modalities can be achieved even in case of partially overlapping volumes without any pre-processing or prior segmentation. Following the suggestion of Viola et al. [12], using the derivative of mutual information and a stochastic approximation our registration procedure became more robust and faster for the 2D case. Concerning the 3D case the expected speed up should be even more dramatic. For the consecutive inspection of our registration results we implemented a tool for superimposing the registered and resampled image data in 2D and 3D. Since there is no single optimal way for fusion there are different tools available which allow the user to select the most appropriate. In order not to compromise any information inherent in the diverse and complex image data we mainly use magic lenses which may be manipulated interactively in size and position. Additionally, the user may navigate with three orthogonal and interacting slices in the volume data.

1. INTRODUCTION

Registration and fusion of tomographic image data from different modalities are important procedures of advanced computer aided visualization in the medical field. In many clinical situations in neuroradiology there are different image data sets available which show either anatomical reference structures, pathological lesions, or functional activity. Depending on the clinical requirements, it is often not sufficient to consider anatomy, morphology, or functional information derived from MR, CT or PET separately but to superimpose images of different modalities and to inspect all the available information simultaneously. Due to different resolution, position, and orientation of the image data and various system inherent sources of distortion a more or less complex transformation must be found which transfers the coordinate system of the floating data set to the coordinate system of the fixed reference data set. In this context registration tries to find the exact geometrical relationship by optimizing a specific measure depending on the applied registration method.

Registration algorithms in medical imaging are mainly divided into two groups. Approaches of the first group are based on artificial, extrinsic image properties mainly employing various stereo tactic frames or skin markers. Since it must be known prior to any imaging procedure whether registration with another modality should be performed these approaches are also classified as non-retrospective. Usually, registration based on stereo tactic frames is very accurate but also very inconvenient for the patient. On the contrary, approaches of the second group try to use image data inherent information exclusively and may therefore be classified as retrospective. There are methods based on anatomical point landmarks, corresponding structures and voxel similarity. Point based procedures are usually very labor intensive with their accuracy depending on the ability of the user to exactly identify corresponding points in different modalities. Surface-based approaches try to align corresponding surfaces delineated in a separate preprocessing step. Since segmentation of these surfaces in images of different modalities is a very crucial task their accuracy is highly data dependent. Within functional data such as PET it is more or less impossible to identify any surface accurately. Finally, there are various voxel based approaches which try to optimize a functional measuring the similarity of all geometrically corresponding voxel pairs. These registration methods seem to be more desirable because the information used for geometrical alignment is not restricted to any specific features. Usually, no pre-processing step is necessary and the registration is not limited by segmentation accuracy. For a more elaborate overview of the different approaches and registration methods consult the reviews of Brown [3], Maurer et al. [4] and van den Elsen et al. [5].

In the first part of this paper we present the applied voxel similarity method based on mutual information. In the second part we present our 2D and pseudo-3D tool for the combined visualization and inspection of the registration results. Finally, we show some of our registration results.

2. REGISTRATION BASED ON MUTUAL INFORMATION

2.1. Overview

Several approaches based on voxel similarity measures have been proposed for the registration of medical image data. Having applied correlation in connection with a surface based approach in scale space [6] the use of joint entropy has been extensively investigated [7, 8, 9, 10]. However, this hasn't proved to be a robust measure of registration since there were major difficulties concerning partial overlap. Collignon et al. [1] and Viola et al. [11, 12] then proposed a much more general information theoretic approach which uses mutual information [13] in order to correctly describe the dispersion in the 2D histogram of two random variables measuring their statistical dependence or the amount of information that one variable contains about the other. With two images of the same subject but of different modalities being superimposed this method states that they are correctly aligned if mutual information of geometrically corresponding grey values is maximal. Since no assumptions are made about the two signals this method is not restricted to specific modalities and does not require the extraction of features in a pre-processing step.

2.2. Method

Neglecting the specific nature of medical image data corresponding grey values of two superimposed images are interpreted as random variables and may simply be read out in an arbitrary order. Then the process of registration can be mapped to the *source - channel - drain* model of information theory with the channel representing the registration transformation T. The image F represented by m different signals $\{f_1, f_2, ..., f_m\}$ of marginal probability distribution $p_F(f)$ is referred to as source or floating image. Its grey values are evaluated at positions x of the reference image R by linear interpolation. This reference image is represented by n different signals $\{r_1, r_2, ..., r_n\}$ of marginal probability distribution $p_R(r)$. After transformation the joint probability distribution of the two signals F and R at corresponding locations x is denoted $p_{RF}(r(x), f(T(x)))$. For convenience we further on use r and $f(T)$ instead of $r(x)$ and $f(T(x))$.

In case there is an exact mapping between the floating image and the reference image (i.e. the signals are identical) the joint probability distribution satisfies

$$p_{RF}(r, f(T)) = p_R(r) = p_F(f(T)) \quad . \tag{1}$$

If the two signals are statistically independent, however, the joint probability distribution satisfies

$$p_{RF}(r, f(T)) = p_R(r) \cdot p_F(f(T)) \quad . \tag{2}$$

Mutual information of R and F is then defined by summing for all grey value pairs $(r, f(T))$ at corresponding positions

$$I_{R,F} = \sum_{(r,f(T))} p(r, f(T)) \log_2 \left(\frac{p(r, f(T))}{p(r)p(f(T))} \right) \quad . \tag{3}$$

It measures the distance between the joint probability distribution $p_{RF}(r, f(T))$ and the distribution $p_R(r) \cdot p_F(f(T))$ of two completely independent signals.

In terms of entropy mutual information is

$$\begin{aligned} I_{R,F} &= H(r) + H(f(T)) - H(r, f(T)) \\ &= H(f(T)) - H(f(T)|r) \\ &= H(r) - H(r|f(T)) \end{aligned} \tag{4}$$

with the entropies $H(r)$ and $H(f(T))$ being a measure of the uncertainty of the random variable r and $f(T)$ respectively. Knowing variable $f(T)$ the conditional entropy $H(r|f(T))$ measures the randomness of r which is left. Therefore, mutual information is the amount of information that $f(T)$ contains about r. Therefore, the registration solution is obtained if mutual information $I_{R,F}$ reaches a maximum for a given set of transformation parameters:

$$I(r, f(T_{opt})) = max_T \left[I(r, f(T)) \right] \quad . \tag{5}$$

If the marginal probability distributions $p_R(r)$ and $p_F(f(T))$ are both independent of the set of transformation parameter T the maximization of mutual information $I_{R,F}$ reduces to the minimization of the joint entropy $H(r, f(T))$. Considering that one image is completely contained in the other the respective marginal probability density becomes independent of the set of transformation parameters T and the optimization of $I_{R,F}$ reduces to the minimization of $H(r|f(T))$ and $H(f(T)|r)$ respectively. Since only the overlapping parts of the superimposed images are evaluated the entropies $H(r)$ and $H(f(T))$ will change depending on the set of transformation parameters T. Mutual information copes with this situation since maximizing the entropies $H(r)$ and $H(f(T))$ tries to attract complex parts of the images whereas minimizing joint entropy $H(r, f(T))$ reduces the dispersion of significant clusters in the 2D histogram.

2.3. Stochastic Derivative Estimation

In the following section an algorithm is described which finds the optimal registration transformation T_{opt} by fast estimation of the derivative of mutual information. This requires the calcualtion of the probability distributions of the random variables r and $f(T)$ which are not directly available. For every iteration they are estimated of a sample A of size $\|A\|$ that holds randomly drawn grid locations of the reference image. The estimation of the densities

$$p(z) \approx \frac{1}{\|A\|} \sum_{x_i \in A} R(z - z_i) \tag{6}$$

with z representing a vector of grey values is then performed with a technique called *Parzen windowing* [14] with $R(z - z_i)$ representing a windowing function. Following the suggestion of Viola et al. [12] a *Gaussian function* G_ψ with variance ψ is used for the 1D case and a diagonal covariance matrix is used for the 2D case. The entropy

$$H(z) \approx -\frac{1}{\|B\|} \sum_{x_i \in B} \ln \left(\frac{1}{\|A\|} \sum_{x_j \in A} G_\psi(z_i - z_j) \right) \tag{7}$$

of one of the images is then approximated with another sample B of size $\|B\|$. In order to find the optimal registration transformation T_{opt} a gradient ascent method is used which globally maximizes the mutual information $I_{R,F}$ of the images R and F. Within this approach the gradient is directly obtained of the derivative of mutual information according to

$$\begin{aligned} &\frac{d}{dT} I(r, f(T)) = \\ &\quad \frac{d}{dT} H(r) + \frac{d}{dT} H(f(T)) - \frac{d}{dT} H(r, f(T)) \ . \end{aligned} \tag{8}$$

Substituting for H equation (7) the approximation of the derivative of mutual information is then given by

$$\begin{aligned} \frac{d}{dT} I(r, f(T)) \approx \frac{1}{\|B\|} \sum_{x_i \in B} \sum_{x_j \in A} (f_i - f_j)^T \cdot \\ \cdot \left(\frac{W_f(f_i, f_j)}{\psi_f} - \frac{W_w(w_i, w_j)}{\{\psi_w\}_{2,2}} \right) \frac{d}{dT}(f_i - f_j) \end{aligned} \tag{9}$$

with the grey values $r_i = r(x_i)$, $f_i = f(T(x_i))$, and $w_i = (r(x_i), f(T(x_i)))$ at corresponding locations x_i of the reference image. The variances ψ_f and ψ_w have been estimated empirically and have been kept unchanged during the registration process. The weights of the 1D case in equation (9) are obtained according to

$$W_f(f_i, f_j) = \frac{G_{\psi_f}(f_i - f_j)}{\sum_{x_k \in A} G_{\psi_f}(f_i - f_k)} \ . \tag{10}$$

For the 2D case the weights $W_w(w_i, w_j)$ are obtained in a similar way to equation (10).

Finally, depending on the applied type of transformation the gradient of the grey values within an image must be determined. In case of a linear transformation T this results to

$$\frac{d}{dT} f(T(x)) = \nabla f(T(x)) \cdot x^T \ . \tag{11}$$

In this way the gradient of mutual information is approximated by a noisy derivative. In order to achieve convergence of the algorithm the computation of $\frac{d}{dT} I(r, f(T))$ must be iterated with different samples A and B for every iteration according to the following pseudo code:

```
for i from 1 to number_of_iterations
{
    A  :=  sample of size ||A||;
    B  :=  sample of size ||B||;
    T  :=  T + λ · (d/dT) I;
}
```

The factor λ is called *learning rate* and must be adjusted for every iteration. In the optimal case it converges to 0 while reaching the global maximum of mutual information.

2.4. Implementation Issues

Image gradients: Equation (11) shows the derivatives for a linear transformations T. In the case of translations the derivative is easily calculated by $\frac{d}{dT} f(T) = \nabla f(T)$. Contrarily general rotations represented by *quaternions* involve extensive calculations. In a more convenient way rotating a point q by a small angle θ is linearly approximated by a cross product according to

$$Rot_{\theta,r}(q) \approx q + (q \times (\theta \cdot r)) \ . \tag{12}$$

For 3D images the combination of translations (x, y, z) and small angle rotations around a fixed axis (x, y, z) leads to a six dimensional derivative in case of the rigid body assumption which is applicable in case of head images.

In order to obtaining $\nabla f(T)$ on a discrete grid *forward differences* were used in the beginning. However, *symmetric differences* are more accurate in a mathematical sense and produce better results.

Randomly selected sample sets: The two sample sets A and B have to be disjoint and should be randomly drawn. They can be obtained by a random number generator which, however, causes a time consuming processing for disjoint samples. Instead a precalculated and static permutation has been used. Shifting this permutation produces pseudo random numbers assuring disjointness of the samples A and B.

Problems and considerations: It was very difficult to decide on the size of A and B. Viola er al. [12] proposed $\|A\| = \|B\| = 50$. We found out that this is sufficient for mono modal registration (CT-CT, MR-MR). However, for multi modal registration (CT-MR) best and fastes results have been achieved with $\|A\| = \|B\| = 100$.

Another problem concerns the estimation of appropriate variances and covariances for the windowing functions in equation (7). We found out that they strongly depend on the size of A and B and the modality of the images. We found out that the variances should be decreased if the size of A and B is increased. However, the analytic dependence of sample size, variance, and image modality requires further investigatation.

As can be seen in formula (9) the time complexity of the algorithm depends cubic on the sample size $N = \|A\| + \|B\|$. Therefore, the runtime behavior of the algorithm is $O(N^3)$. However, we reduced it to $O(N^2)$ by calculating the denominator of equation (10) only once since it didn't vary much. Producing results which have been as good as beforehand we decided on using this pure empirical approach.

3. FUSION OF REGISTERED IMAGES

Once the registration procedure was successful the registration transformation T_{opt} is applied and the floating data set is re-sliced according to the reference data set. This results in both data sets having the same volume and the same voxel dimensions. Since both data sets may then be regarded as a single combined image with two attributes at each voxel fusion is a logical consequence. However, simultaneous visualization of complementary parameters is a difficult if not impossible task if no information should be lost.

Considering the increasing amount of medical images visualization of only significant information after registration should make perception and interpretation of the available data much easier and may therefore effectively assist diagnosis and therapy planning. However, if only relevant information is shown it is even more difficult to decide what information should be preserved and what rejected. Since the registration result is normally presented in a visual form to the diagnostician a single optimal way of fusing registered images may not exist. Therefore, it seems to be more realistic to provide fusion tools which allow to interactively adjust the most appropriate representation of the data or to clearly visualize every attribute separately in partitioned displays.

In order to visualize and inspect the results of our registration procedure a simple but effective tool for the visual fusion of registered 2D and 3D image data has been implemented. As can be seen in Figure (2) it offers three displays showing orthogonal slices of the image data with one image in the *background* and the other image in a window. The window or *magic lense* can be move or resized interactively and simultaneously with the mouse in all three displays, and therefore allows intuitive comparison of the registered images without compromising any information. Further on, a pointer can be moved with the mouse in order to navigate through the volume changing the point of intersection of the orthogonal slices.

Additional features for the visualization and merging of the images are orthogonal/horizontal stripes, chess board, zoom-in/out, MIP and α-blending (including a user defined threshold option). Finally, the images in the background window and the magic lense can be exchanged.

4. RESULTS AND DISCUSSION

In order to evaluate the registration approach based on mutual information with stochastic gradient estimation, and in order to check its critical parameters the algorithm has been implemented for the 2D case. Test slices have been taken from a 3D registration performed with the histogram approach which already proved to produce excellent results. Despite of a considerable sub-sampling of the data the new method produced parameters with negligible differences in comparison to the histogram method with the registration solution being found within 30-120 seconds on a 195MHz SGI Indigo^2 R10000. Since the size of the samples A and B will not essentially increase for the 3D case a tremendous speed up is expected compared to the histogram method. Figure 1 shows the initial situation of a 2D CT-MR registration (top/middle) and the superimposed result images (bottom). Additionally, Figure 2 shows our fusion tool used for the visualization of registered 2D and 3D image data presenting the checker board feature at the top and the magic lenses at the bottom.

5. CONCLUSION

Registration based on mutual information and stochastic gradient estimation as presented by Viola et al. [12] was successfully applied to 2D MR-CT data. Critical parameters like the variances and covariances of the windowing functions and the size of the samples have been investigated. Due to the robustness of the algorithm and due to the expected speed up our implementation will be extended to 3D in the near future including first tests with non-rigid transformations.

6. REFERENCES

[1] A. Collignon, D. Vandermeulen, P.Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. *Kluwen Acad. Publ's: Computational Imaging and Vision*, 3:263–274, 1995.

[2] P. Viola. *Alignment by Maximization of Mutual Information.* PhD thesis, M.I.T., 1995.

[3] L.G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, Dez 1992.

[4] C.R.Maurer and J.M. Fitzpatrick. A review of medical image registration. In *Interactive Image-Guided Surgery*, pages 17–44, Maciunas, R.J. (Ed), Park Ridge, IL, American Association of Neurological Surgeons, 1993.

[5] P.A. van den Elsen, E.-J.D. Pol, and M.A. Viergever. Medical Image Matching - A Review with Classification. *IEEE Eng. in Medicine and Biol.*, pages 26–38, March 1993.

[6] P.A. van den Elsen. *Multimodality Matching of Brain Images.* PhD thesis, Utrecht University Thesis, 1993.

[7] D.L.G. Hill, C. Studholm, and D.J. Hawkes. Voxel Similarity Measure for Automated Image Registration. In Proc. SPIE, editor, *Visualization in Biomedical Computing*, volume 2359, Bellingham, WA, Rochester, Oct 1994.

[8] A. Collignon, D. Vandermeulen, P.Suetens, and G. Marchal. 3d multi-modality medical image registration using feature space clustering. *Lecture Notes in Computer Science*, 905:195–204, 1995.

[9] C. Studholme, D. Hill, and D. Hawkes. Automated 3D Registration of Truncated MR and CT Images of the Head. In P. Pycock, editor, *Proc. of the British Maschine Vision Conf. (BMVC)*, pages 27–36. Britisch Maschine Vision Association, 1995.

[10] C. Studholme, D. Hill, and D. Hawkes. Multiresolution Voxel Similarity Measure for MR-PET Registration. In Y. Bizais, editor, *Proc. of the Inf. Proc. in Med. Imaging Conf.* Kluwer Academic Publisher, 1995.

[11] P. Viola and W.M. Wells. Alignment by Maximization of Mutual Information. In *Proc. Vth Int. Conf. Computer Vision*, pages 16–23, Cambridge, MA, 1995.

[12] W.M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis, Oxford University Press*, 1(1), March 1996.

[13] M. Bossert. *Kanalcodierung.* Teubner Verlag, Stuttgart, 1992.

[14] R. Duda and P. Hart. *Pattern Classification and Scene Analysis.* John Wiley and Sons, 1993.

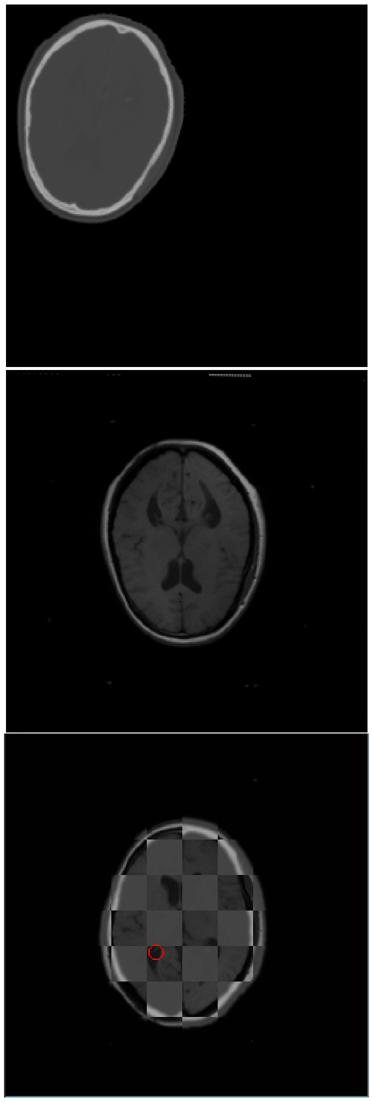

Figure 1: Initial position of CT (top) and MR (middle) images and fused registration result (bottom)

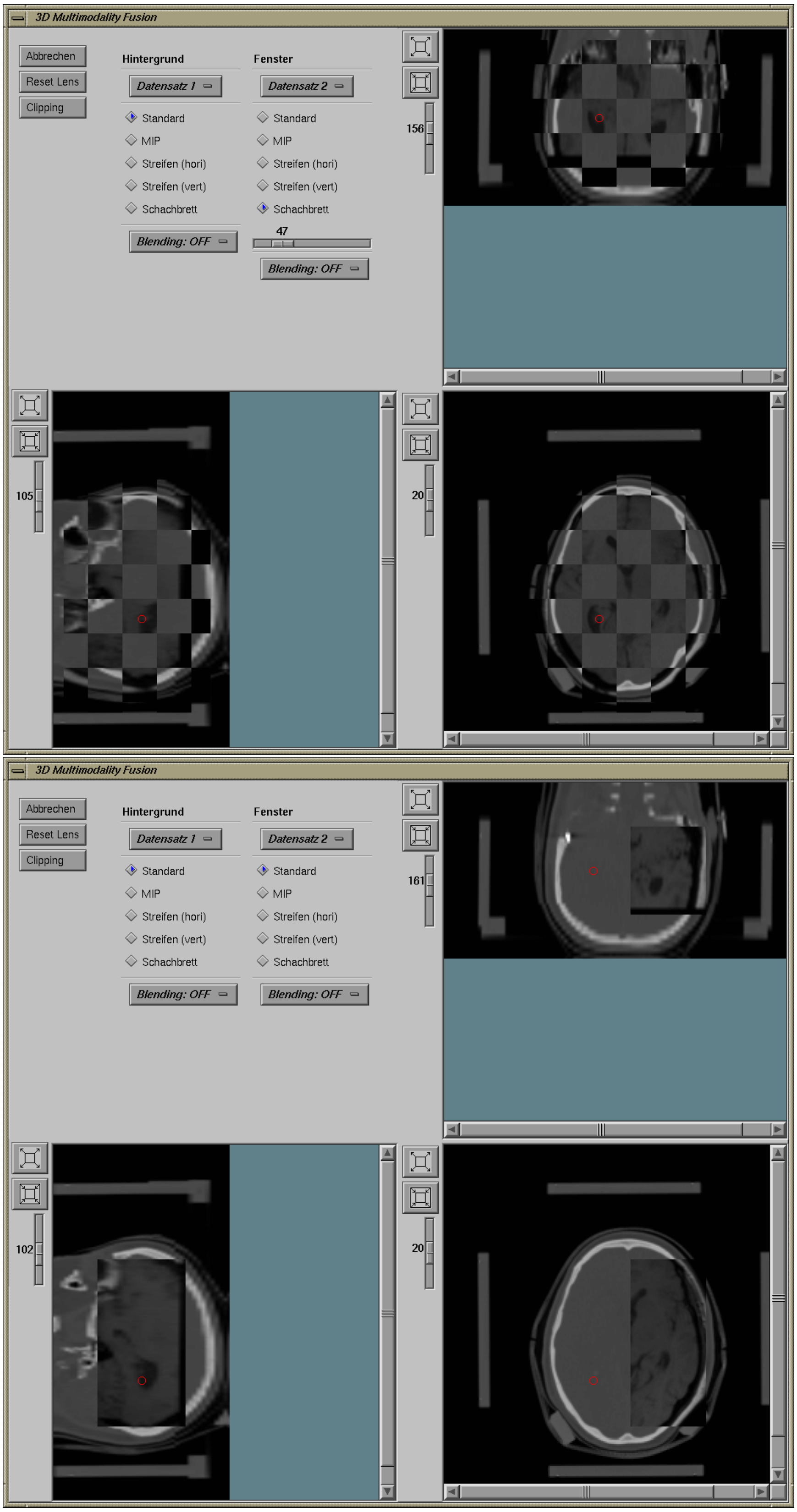

Figure 2: Fusion tool for the visualization of registered image data showing magic lense and checkerboard