Adaptation of Mouse Brain Gene Expression Data for further Statistical Parametrical Mapping Analysis

Anton Osokin[(1)], Alexey Lebedev[(2)], Dmitry Vetrov[(1)], Vladimir Galatenko[(2)], Dmitry Kropotov[(3)], Alexandr Nedzved[(4)], Anton Konushin[(1)], Konstantin Anokhin[(5)] *

(1) Department of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia
osokin.anton@gmail.com, vetrovd@yandex.ru, ktosh@graphics.cs.msu.ru http://cs.msu.ru

(2) Department of Mechanics and Mathematics, Moscow State University, Moscow, Russia
alleb@list.ru, vgalat@castle.nmd.msu.ru, http://www.math.msu.ru

(3) Dorodnicyn Computing Center of the Russian Academy of Sciences
dkropotov@yandex.ru, http://www.ccas.ru

(4) United Institute of Informatics Problems, National Academy of Sciences of Belarus
abelotser@newman.bas-net.by, http://http://uiip.bas-net.by

(5) P.K. Anokhin Institute of Normal Physiology. E-mail:
k.anokhin@gmail.com, http://www.aha.ru/ niinf

Abstract

The paper describes a method for fully automatic 3D-reconstruction of mouse brain voxel model from a sequence of coronal 2D slices for statistical analysis of gene expression. Two images of each brain slice with different stains are used. The first stain highlights the histology of brain, which is used for slice matching. The second stain highlights the level of gene expression. The algorithm proceeds as follows. First, images are preprocessed to suppress image noise and equalize image brightness. Second we estimate the level of gene expression in each slice using the second stain. Then we construct 3D-model of the brain using the first stain. To do this all images are aligned via rigid-body transformations. After alignment neighboring slices are matched by estimation of non-linear deformations. As the distance between slices is significantly larger then image resolution we add intermediate virtual slices using morphing algorithm. Gene expression level is interpolated in identical way. The obtained 3D-model with the information about gene expression can be used for gene expression analysis via Statistical Parametric Mapping (SPM) package. The proposed method for 3D-reconstruction has been tested on images from Allen Brain Atlas, which is available in electronic form.

Keywords: *3D-Reconstruction, neuroimaging, image transformations, morphing, elastic deformations, image registration, B-splines.*

1. INTRODUCTION

The problem of gene expression analysis using only images of brain slices is very important in modern brain research [1]. Neurobiologists are now able to measure the activity of selected gene in every brain cell. This is usually done in vitro, i.e. on dead species. The extracted brain is frozen and then cut into slices. Each slice is double-stained by Nissl method to highlight histology and by special stain which reveals the neurons with expression of corresponding genes.

The further automatic analysis of gene expression is significantly more complicated problem, compared to analysis of brain activity, measured by other neuroimaging techniques, like fMRI or PET. For fMRI data a well-known Statistical Parametric Mapping (SPM) framework is widely used[2]. SPM uses 3D voxel model as input data and provides tools for model reconstruction from series of slice images. However, SPM package cannot be directly applied for gene analysis problem. First, due to technological aspects of brain cutting procedure brain slices differ in their actual size, shape and orientation, and SPM fails to correctly reconstruct model from such kind of data. Second, statistical analysis in SPM package relies on spatial coherency of brain activity, which is the reason why it uses uniform voxel models. fMRI data is acquired in fairly low resolution (128*128 or 256*256 images are a common case) and it is easy to acquire similar number of fMRI images to ensure that voxel model has same resolution in all dimensions. But for gene expression analysis image resolution should be significantly higher, and it is impossible to obtain corresponding number of slices for many technological reasons. Thus it is necessary to reconstruct intermediate layers from existing slices.

In this paper we propose a novel method for voxel model reconstruction from the set of brain slice images that allows further analysis by SPM package. Our method is fully automatic and consists of following steps. First, images are preprocessed to suppress image noise and equalize image brightness. Then all images are aligned via rigid-body transformations. Consecutive slices are matched by estimation of non-linear deformations between layers. Intermediate layers are then interpolated using estimated non-linear transformations. Finally, data is converted to the voxel models and saved in the NIFTI-1 file format [3] that is standard for the brain studies.

We have tested our method 3D-reconstruction on images from Allen Brain Atlas [4] and showed that it leads to comparable results with the ones in AGEA project where 5 times more slices are used for 3D-reconstruction.

The rest of paper is organized as follows. In section 2 we describe the preprocessing of brain slice images with histological and gene expression stains. Section 3 describes how gene expression level is computed. In section 4 image alignment and intra-slice interpolation method is described. Section 5 contains some experimental results. Some conclusions are given in the last section.

2. BRAIN SEGMENTATION AND ILLUMINATION CORRECTION

On experimental images of mouse brain slices noise level is very high due to uneven amount of stain in different part of the slice. This results in abrupt deformations of brain slice contour that de-

*This work was supported by Russian Foundation for Basic Research, projects Nos. 08-01-00405, 08-01-90016 and 09-04-12215.

grade alignment and matching. To suppress such defects we apply segmentation algorithm based on graph cuts with prior assumptions on slice shape. [5, 6, 7].

Uneven illumination of brain images is also a problem. For different brain images contrast is different and even within one separate image there are areas with different illumination levels. For illumination correction we apply Single Scale Retinex method [8]. First, Gauss filter with large radius is applied to the image:

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, -R \le x,y \le R.$$

We used $\sigma = 20, R = 100$. This operation gives an illumination map of the image I_{map}.

Afterwards the initial image is divided by obtained filtered image

$$I_{new}(i,j) = \frac{I(i,j)}{I_{map}(i,j)+1}.$$

This operation provides both equal local illumination within each image and equal illumination of different atlas images.

Figure 1 and figure 2 illustrate this procedure.

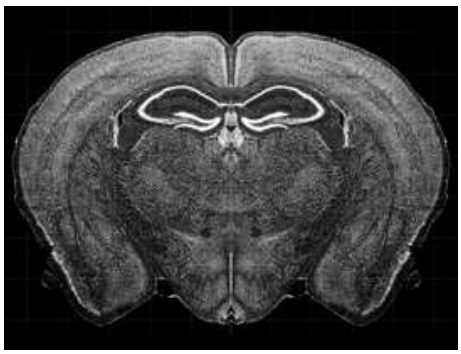

Figure 1: Histological brain image after Graph Cut algorithm without illumination correction.

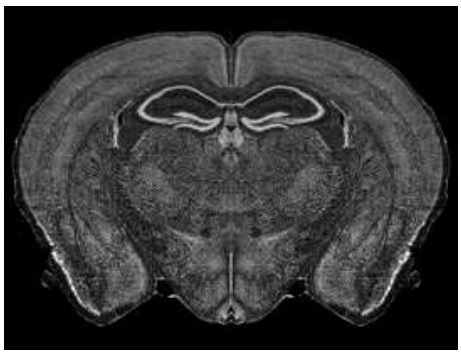

Figure 2: Histological brain image after illumination correction.

3. EXPRESSION DETECTION AND EVALUATION

To measure the gene expression in stained brain slice its image is acquired with very high resolution so that each stained nucleus is visible. But such high resolution make further processing and analysis computationally impractical and it should be reduced. Straightforward resolution decrease will smooth image and introduce errors in gene expression measurement. So we need a specific procedure for rescaling of such images.

For this purpose an expression level is evaluated for each image zone that corresponds to one pixel after the resolution reduction. First we evaluate the expression level for each pixel of the initial image, then apply a standard Gauss filter with radius close to the new pixel size (this also leads to smoothing required for the subsequent statistical analysis), and finally evaluate the expression level for each new pixel by simple summation and rescaling. Note that in following versions we plan to replace the application of Gauss filter by non-uniform smoothing method that takes account of anatomic structures.

The results of expression level detection and evaluation can be corrected manually if required by changing the default thresholds and other parameters (globally or locally, in a certain image fragment) and/or by setting the expression level in certain image points manually. However the designed procedure of expression detection and evaluation is fully automated and normally the manual correction is not required.

The most critical part of this step is the evaluation the expression level for each pixel of the initial image. In current implementation the fact of expression presence in a point is not simply binary, but is represented by a real number from 0 (no expression) up to 1 (maximum expression level). Intermediate values represent the relative expression level. Such approach is natural from the biological viewpoint, it also leads to the correctness of the computations (where correctness is understood in a standard mathematical sense) and hence eliminates the errors connected with close-to-threshold values.

The procedure of expression evaluation can be shortly described as follows. The image is first transformed from RGB to one-parametric color palette (in fact it is an analogue of the grey-scale palette, but different coefficients can be used for the transformation). Similarly to [9] two different one-parametric color schemes are used in order to reduce errors. For each color scheme the value of a sigmoid function (i.e., a smooth function that equals 0 for arguments less than the lower threshold and equals 1 for arguments greater than the upper threshold) is computed at each pixel: the color of the pixel in the one-parametric color palette is used as an argument, and the parameters of the sigmoid function depend on the distance from the brain border and the mean color value in the large neighborhood. The expression activity at each pixel is set to the product of the values of these sigmoid functions. Figures 3, 4 show the simplified result of expression detection.

4. 3D MODEL CONSTRUCTION

4.1 Image alignment

In fMRI imaging the alignment transformations between images are assumed to be either affine or rigid-body in 3D or 2D space [10]. In our case slices are already horizontally centered and aligned w.r.t symmetry line during imaging process. But cutting procedure introduces small deformations in vertical direction, which are different for each slice. So we limit the set of alignment transforms by vertical shifts and stretches.

Image alignment proceeds as follows. First, we search for the smallest surrounding rectangle of each slice. Then we consider rectangle border to be a function of slice number. Figure 5 shows example of top and bottom borders of brain rectangles without alignment. These functions are not smooth enough for 3D reconstruction. To smooth these functions we apply Savitzky-Golay filter.

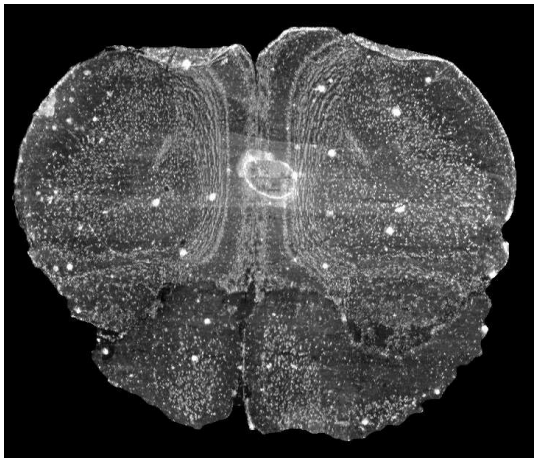

Figure 3: Initial image of a brain region.

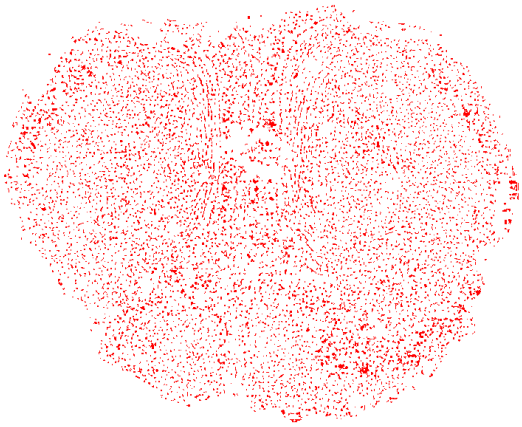

Figure 4: Image of the brain region where pixels with expression level greater than threshold (0.5) are marked.

The idea of Savitzky-Golay filtering is to find filter coefficients that preserve higher moments. Equivalently, the idea is to approximate the underlying function within the moving window not by a constant (whose estimate is the average), but by a polynomial of higher order (we used order 5 in our approach). For each point we least-squares fit a polynomial to the points in the moving window (we used window width 15), and then set the new value to be the value of that polynomial at the same position.

Savitzky-Golay filters can be thought of as a generalized moving average. Their coefficients are chosen this way to preserve higher moments in the data, thus reducing distortion of essential features of data like peak heights and line widths in a spectrum, while the suppression of random noise is improved. Figure 6 shows example of top and bottom borders of brain rectangles after alignment.

If we are interested in any specific section of mouse brain we can make additional alignment in appropriate plane. Such alignment makes specific section smoother but the whole model becomes less smooth. So in general we don't use specific plane alignment for 3D-model reconstruction.

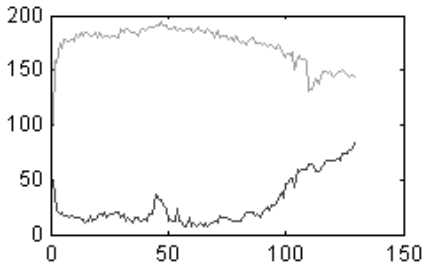

Figure 5: Not aligned top (U) and bottom (D) borders of mouse brain.

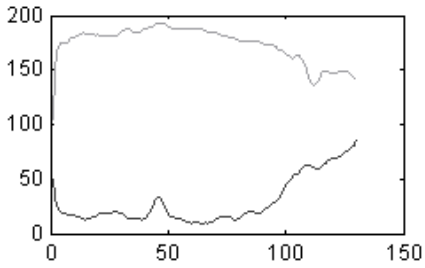

Figure 6: Aligned top (U) and bottom (D) borders of mouse brain.

4.2 Non-linear transformations

To obtain the voxel model with uniform resolution we need to fill the gaps between available slices by synthesizing of intermediate layers. For this task we estimate non-linear deformation between consecutive slices.

There exist many methods for non-linear deformation estimation: parametric model of the deformation [10, 11, 12, 13], hierarchical models [14, 15], nonparametric local methods [16], inclosed dynamic programming [17], optical flows [18]. In this paper we follow the papers [11, 12] and use the approach based on parametric model of deformation based on B-spline basis functions. Choice of B-splines as basis functions provides good quality of deformation and high speed of calculation because of limited number of basis functions.

A brain 3D model is a function:

$$F : \mathbb{R}^3 \to [0, 1].$$

From slices we know F values only at some discrete points. In slice plane expansion of discrete function to its continuous version is a weighted sum of surrounding discrete point colors. Expansion in other plane can be done in the same way (weighted sum of neighboring slices). However, this simple solution makes a 3D model not smooth enough. A better solution can be obtained using non-linear image deformations.

The input images are given as two 2-dimensional discrete functions:

$$f_1, f_2 : I \subset \mathbb{Z}^2 \to [0, 1].$$

Here I is a 2-dimensional discrete interval covering the set of all pixels in the image. Function values stand for intensities of corresponding pixels.

We denote continuous expansions of two images as f_1^c, f_2^c.

Our goal is to find a deformation of the first image to the second one in the following way:

$$f_1^c(g(x,y)) \approx f_2(x,y).$$

Here $g(x,y) : \mathbb{R}^2 \to \mathbb{R}^2$ is a deformation (correspondence) function between pixels.

We measure the difference between images by SSD (sum of squared deviations) criterion:

$$E = \sum_{(i,j)\in I} (f_1^c(g(i,j)) - f_2(i,j))^2.$$

So the problem is to minimize E with respect to deformation function g.

We consider deformation function as a linear combination of some basis functions:

$$g(x,y) = \sum_{k\in K} \overrightarrow{c_k} b_k(x,y).$$

Here K is a set of basis function indexes.

Family of deformation functions 4.2 transforms optimization problem in functional space into finite-dimensional optimization problem.

We use uniformly spaced cubic B-splines as basis functions.

A B-spline β_r of degree r is recursively defined as

$$\beta_r = \beta_{r-1} * \beta_0, r > 0.$$

β_0 is a characteristic function of $[-0.5, 0.5]$, $*$ is convolution operator.

Specifically, cubic B-spline is the following function:

$$\beta_3(x) = \begin{cases} 2/3 - (1-|x|/2)x^2, 0 < |x| \le 1, \\ (2-|x|)^3/6, 1 < |x| < 2, \\ 0, |x| \ge 2. \end{cases}$$

So we are looking for the deformation function in the family:

$$g(x,y) = \sum_{(k_x,k_y)\in K} \beta_3(x/h_x - k_x)\beta_3(y/h_y - k_y).$$

Centers of B-spline functions are placed on the regular grid $(k_x h_x, k_y h_y)$. Working with uniform splines is significantly faster with respect to nonuniform splines. In order to get complete control over g, we put some spline knots outside the image.

Finally the problem is to optimize SSD criteria E w.r.t. set of parameters c. Here we use gradient descent algorithm with feedback step size adjustment. In this algorithm parameter update rule is $\Delta c = -\mu \nabla_c E(c)$. After a successful step μ is multiplied by some value $\mu_f > 1$, otherwise it is divided by some other value $\mu_f^* > 1$.

An example of deformation field obtained from B-spline basis functions for a pair of neighboring slices from Allen Atlas is shown in figure 7.

Since we have deformation of the first image to the second one and vice versa, we can fill gaps between atlas slices with weighted sum of deformed neighboring slices:

$$F(x,y,z) = (1-\alpha) f_{1,k-1}^{\alpha}(x,y) + \alpha f_{2,k}^{1-\alpha}(x,y).$$

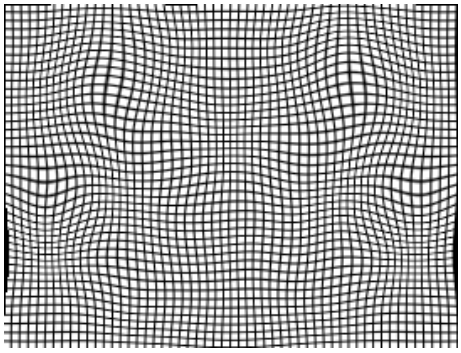

Figure 7: Deformation field for B-spline method. This deformation field is obtained by applying deformation of neighboring slices to regular grid.

Here $\alpha = \frac{z - z_{k-1}}{z_k - z_{k-1}}$, $z_{k-1} \le z < z_k$, z_k is a z-coordinate of slice number k.

$$f_{1,k-1}^{\alpha}(x,y) = f_{k-1}((x,y) + \alpha(g_{k-1}^{k}(x,y) - (x,y))).$$

$$f_{2,k}^{1-\alpha}(x,y) = f_k((x,y) + (1-\alpha)(g_k^{k-1}(x,y) - (x,y))).$$

Here $g_i^j(x,y)$ is a deformation function of slice number i to slice number j.

As a last step, the full set of input and intermediate slices is converted to voxel model and saved in the NIFTI-1 file format [3], which is a standard for the brain studies.

5. EXPERIMENTAL RESULTS

In order to test the proposed method for 3D-model construction we used it on coronal Allen Brain Atlas [19, 4].

The coronal Allen Brain Atlas is a set of full-color, high-resolution coronal digital images (132 images) of mouse brain accompanied by a systematic, hierarchically organized taxonomy of mouse brain structures. The Allen Brain Atlas is obtained from 8-week old C57Bl/6J male mouse brain prepared as unfixed, fresh-frozen tissue.

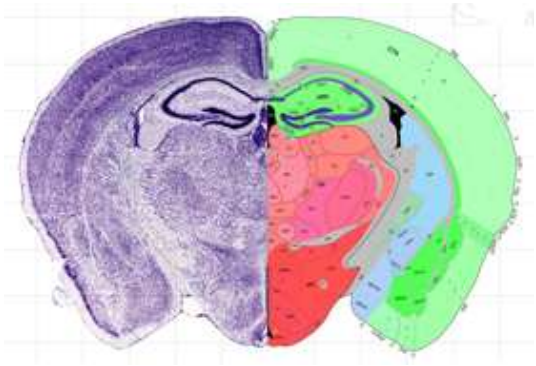

Figure 8: Allen Brain Atlas image.

On figure 8 the left half represents histological structure of one mouse brain slice. The right half shows color annotation of mouse brain structures made by human experts. We have used histological images to reconstruct a 3D voxel model. Afterwards we applied same morphing transformations to create virtual slices of annotated

images to demonstrated that our method correctly preserves the inner brain structures.

Figures 9 and 10 show synthesized histological and color annotated axial slices of model built without illumination correction and slice alignment. Figures 11 and 12 show synthesized histological and color annotated axial slices of model built with both illumination correction and slice alignment, but without non-linear deformations between neighboring slices.

Figures 13 and 14 show synthesized histological and structural axial slices of model built with illumination correction, slice alignment and with non-linear deformations between neighboring slices.

These figures show significant improvement of 3D model quality with each step. 3D model from figures 13 and 14 is much smoother than the previous ones. It should be mentioned that not only brain borders become smoother but also borders of internal structures.

We also compared our method of virtual slices generation with the analogous method used in AGEA project [20]. Figure 15 shows appropriate axial brain view from AGEA project. The quality of view, obtained by our method, is very similar to one of AGEA project, but compared compared to full set of slices without gaps, used in AGEA, we have used only each 5th slice.

The results of the automated gene expression detection and evaluation were checked by human experts and were found to be satisfactory.

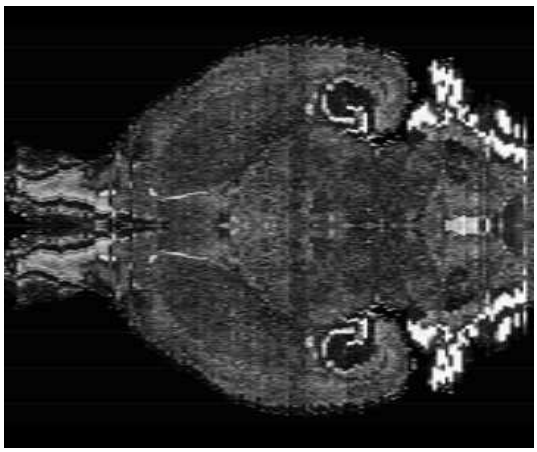

Figure 9: Axial histological view of 3D model without illumination correction, alignment and nonlinear deformations.

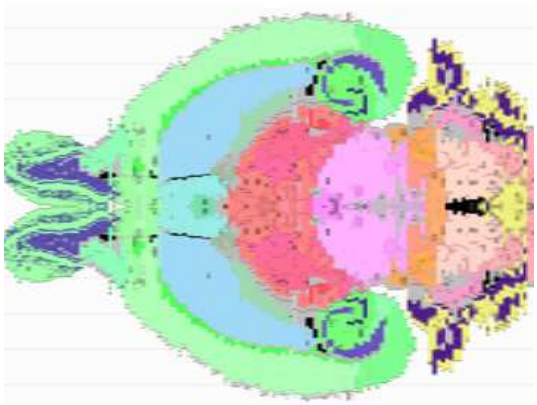

Figure 10: Axial structural view of 3D model without illumination correction, alignment and nonlinear deformations.

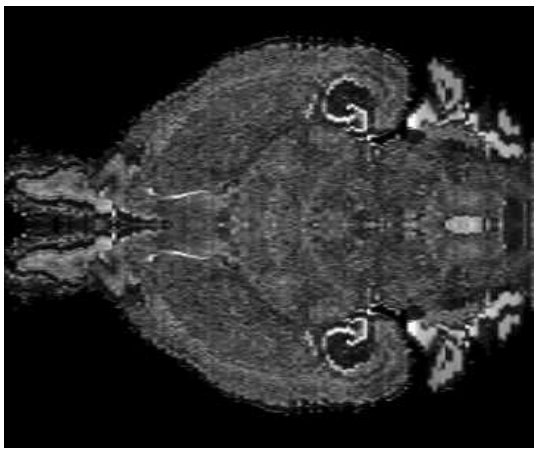

Figure 11: Axial histological view of 3D model with illumination correction and alignment.

6. CONCLUSION

We proposed an algorithm that constructs virtual slices of brain w.r.t. arbitrary section-plane. We have shown that such algorithm allows us to get synthetic images of relatively good quality both with histological and anatomical structure. Such algorithm opens great perspectives for further brain research as it provides the opportunity of discovering the anatomical structures in a single slice of real mouse brain. The procedure of slices' preparation is very time and labor consuming, that is why it is highly desirable to reduce the number of slices obtained from real mouse to minimum (in the limit to one which is of interest for biologists). The slice can be made in non-standard (coronal, sagittal, or axial) section-plane and it should be mapped into 3D-model of atlas brain. Our algorithm allows us to synthesize the image of an atlas brain w.r.t. any section-plane and hence is the key part of future method which will compute the best mapping. When it is done the anatomical structures in real brain slice can be found easily by projecting anatomical structure of atlas brain onto the virtual slice with further performing inverse mapping to adapt it to real brain slice. The algorithm for identifying anatomical structures in arbitrary brain slice is extremely useful for brain research as it allows to understand what structures are responsible for specific genes expression in specific situations.

Note that the quality of statistical analysis (including SPM analysis) can be significantly increased by using the knowledge about anatomic structures. A simple way of structural annotation of a model brain is based on the geometric alignment of this model with the model that integrates histologic data and anatomic (structural) data. Such model was constructed based on Allen Brain Atlas. The alignment can be built, e.g, using standard SPM procedures for geometric processing.

7. REFERENCES

[1] L. Ng, S.D. Pathak, C. Cuan, C. Lau, H. Dong, A. Sodt, C. Dang, B. Avants, P. Yushkevich, J.C. Gee, D. Haynor, E. Lein, A. Jones, and M. Hawlyrycz, "Neuroinformatics

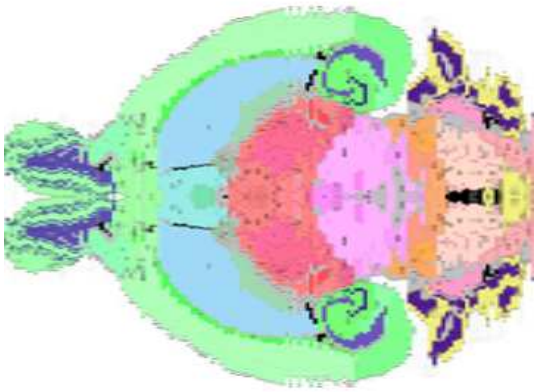

Figure 12: Axial structural view of 3D model with illumination correction and alignment.

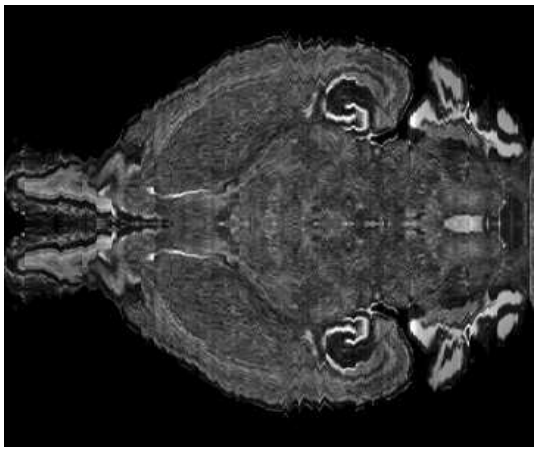

Figure 13: Axial histological view of 3D model with illumination correction, alignment and nonlinear deformations.

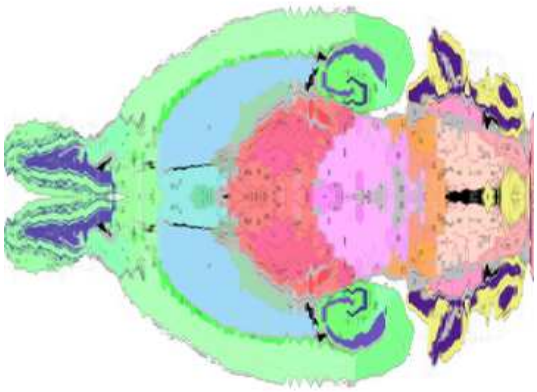

Figure 14: Axial structural view of 3D model with illumination correction, alignment and nonlinear deformations.

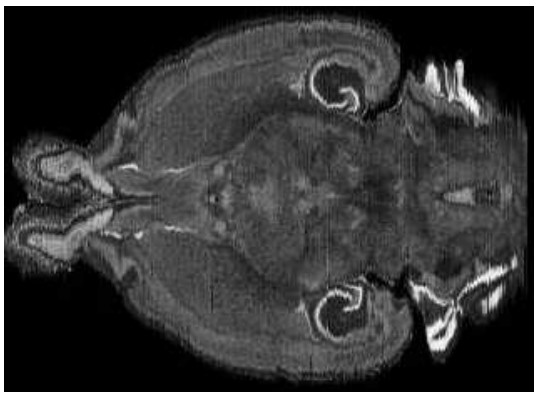

Figure 15: Axial histological view of 3D model of AGEA project.

for genome-wide 3d gene expression mapping in the mouse brain," *IEEE Transactions on Computational Biology and Bioinfomatics*, vol. 4, no. 3, pp. 382–393, 2007.

[2] "Spm8," Internet, Available from: http://www.fil.ion.ucl.ac.uk/spm/, 2009.

[3] "Nifti-1 file format," Internet, Available from: http://nifti.nimh.nih.gov/, 2009.

[4] Seattle (WA): Allen Institute for Brain Science, "Allen mouse brain atlas," Internet, Available from: http://mouse.brain-map.org, 2009.

[5] Y. Boykov, O. Veksler, and R. Zabih, "Efficient approximate energy minimization via graph cuts," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1222–1239, 2001.

[6] V. Kolmogorov and R. Zabih, "Energy functions can be minimized via graph cuts?," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, 2004.

[7] Y. Boykov and V.Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.

[8] Y. Boykov and V. Kolmogorov, "Multi-scale retinex for color image enhancement," *Proceedings of ICIP*, vol. 3, pp. 1003–1006, 1996.

[9] J. P. Carson, T. Ju, C. Thaller, J. Warren, M. Bello, I. Kakadiaris, W. Chiu, and G. Eichele, "Automated characterization of gene expression patterns with an atlas of the mouse brain," *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, vol. 2, pp. 2917–2920, 2004.

[10] R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny, *Human Brain Function*, Academic Press, 2nd edition, 2003.

[11] J. Kybic, P. Thevenaz, A. Nirkko, and M. Unser, "Unwarping of unidirectionally distorted epi images," *IEEE Transactions on Medical Imaging*, vol. 19, no. 2, pp. 80–93, 2000.

[12] J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1427–1442, 2003.

[13] Y. Wu, T. Kanade, C. Li, and J. Cohn, "Image registration using waveletbased motion model," *International Journal of Computer Vision*, vol. 38, pp. 129–152, 2000.

[14] P. Moulin, R. Krishnamurthy, and J. Woods, "Multiscale modeling and estimation of motion fields for video coding," *IEEE Trans. Image Processing*, vol. 6, pp. 1606–1620, 1997.

[15] O. Musse, F. Heitz, and J.-P. Armspach, "Topology preserving deformable image matching using constrained hierarchical parametric models," *IEEE Trans. Med. Imag.*, vol. 10, pp. 1081–1093, 2001.

[16] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching.," *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 1–21, 1989.

[17] T. Ju, J. Warren, J. Carson, M. Bello, I. Kakadiaris, W. Chiu, C. Thaller, and G. Eichele, "3d volume reconstruction of a mouse brain from histological sections using warp filtering," *Journal of Neuroscience Methods*, vol. 156, pp. 84–100, 2006.

[18] J.L. Barron and D.J.Fleet ans S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.

[19] E.S. Lein et al.., "Genome-wide atlas of gene expression in the adult mouse brain," *Nature 445*, pp. 168–176, 2007.

[20] Seattle (WA): Allen Institute for Brain Science, "Agea project," Internet, Available from: http://mouse.brain-map.org/agea/, 2009.

ABOUT THE AUTHORS

Anton Osokin — 4th-year student of Computational Mathematics and Cybernetics department of Moscow State University. E-mail: Osokin.Anton@gmail.com

Lebedev Alexey — 3rd-year student of Mechanics and Mathematics department of Moscow State University. E-mail: alleb@list.ru

Dmitry Vetrov — PhD, researcher in the chair of Mathematical Methods of Forecasting, Computational Mathematics and Cybernetics department of Moscow State University. E-mail: VetrovD@yandex.ru

Galatenko Vladimir — PhD, associate professor of Mechanics and Mathematics department of Moscow State University. E-mail: vgalat@castle.nmd.msu.ru

Dmitry Kropotov — junior researcher in Dorodnicyn Computing Centre of the Russian Academy of Sciences. E-mail: DKropotov@yandex.ru

Alexandr Nedzved — PhD, researcher in United Institute of Informatics Problems, National Academy of Sciences of Belarus. E-mail: abelotser@newman.bas-net.by

Anton Konushin — PhD, researcher in Graphics and Mutlimedia Lab, Computational Mathematics and Cybernetics department of Moscow State University. E-mail: ktosh@graphics.cs.msu.ru

Konstantin Anokhin — doctor of medicine, member-correspondent of the Russian Academy of Medical Sciences, head of systemogenesis department in P.K. Anokhin Institute of Normal Physiology. E-mail: k.anokhin@gmail.com