УДК 004.932

T. S. Mandziy

ACTIVE SHAPE MODELS WITH ADAPTIVE WEIGHTS

Improvement of active shape models dealing with noisy and blured images of objects is developed. Results are tested on a set of radiographic images of welds. Improvement of performance of the proposed modified active shape models for radiographic images compared to conventional ones was shown basing on experimental results comparison.

Удосконалено моделі активних форм для зашумлених зображень та зображень з нечіткими границями об'єктів. Результати протестовано на наборі рентгенографічних зображень зварних швів. Покращання функціонування вдосконалених моделей активних форм порівняно з класичними продемонстровано на порівнянні експериментальних результатів для рентгенографічних зображень.

Representation of images containing objects, whose shape can vary, is a necessary and challenging task. Examples of such objects can be images of human faces or magnetic resonance brain sections. Models called to deal with such type of object images have very wide practical application. They are able to handle various types of tasks, such as object shape modeling, objects tracking, feature localization, image segmentation etc. They also can be used for object state recognition, like recognition of human emotions. Those models are very important for tasks of image segmentation where correct segmentation is impossible without relying on shape information due to its low signal to noise ratio or texture inconsistency. Such model application is very important and successful in medicine and industry for radiographic, magnetic resonance and ultrasonic image segmentation.

There were proposed many models to handle aforementioned tasks. Probably, the first attempt to deal with images of varying objects were active contours [1]. They are able to track changes in shape, but they are shape free models and initially were not able to respond to the particular shapes. There exist some improvements of the model to make it sensitive to the particular user defined shapes [2, 3].

Proposed by T. Cootes Active Shape Models (ASM) (also known as smart snakes) [4] were designed especially for handling shape variations of object images. The ASM is relying on statistical model of shape variation. Shape in this model is represented as a set of landmarks (set of points placed on a statistically significant object image parts). Relative variations of landmarks are constrained by a Point Distribution Models (PDM) captured from a training set of shapes. Matching model to image is made by iterative technique. A new landmark point locations are obtained by nearby search around current landmark point locations, aiming to find the best texture model match, expected at the landmark position, with the image area around current landmarks. After new landmark point locations are found, parameters of a model are adjusted to the best match of these new locations to model generated ones. Since T. Cootes original paper was published there were made a lot of efforts to improve the ASM: double contours ASM [5], ASM with bifurcation contours handling [6], non-linear multi-view ASM [7] etc.

The following reveals mathematical basis of ASM, gives the examples of cases where conventional ASM is failed to show a consistent result, proposes an improvement of the model to deal with those hard cases and compares obtained results to those obtained with conventional ASM.

Active shape model formulation. ***Model building***. ASM is built based on training set of annotated images (images with placed ground truth landmarks) of modeled object. There are two main stages in model building. The first is building statistical model of shape. This step uses only coordinates of landmarks as an input and disregards images. The second is a landmark local texture descriptions building. This step utilizes images and landmark position information.

Active shape models represent an object shape by a fixed size set of key points that correspond to particular previously selected interest regions of the object (e.g. those points are usually placed along the contours of significant object parts, such as nose, mouth, eyes etc., if images of human faces are considered).

Point distribution model (PDM) [4] is a statistical tool for modeling variations of object shape. The set of all possible shapes is assumed to form a Gaussian distribution around some mean point $\overline{x} = (\overline{x}_1, ..., \overline{x}_n, \overline{y}_1, ...\overline{y}_n)$ in shape space. PDM is inferred, based on training set composed of possible shape variations examples. After the training data are collected all training shapes must be aligned in minimum of mean squared error sense (1):

$$\min \sum_i (x_i - \overline{x})(x_i - \overline{x})^T , \qquad (1)$$

where $\overline{x}$ is the mean shape and x_i is an i^{th} shape from training set. Interested readers are referred to [4] for introduction with technique of aligning two similar shapes by removing relative translation, rotation and scale between them. To build a compact shape model (i.e. to build model with possible smallest model parameters set) principal component analysis (PCA) is applied. After mean shape $\overline{x}$ and PCA transform is known arbitrary shape $\mathbf{x}$ similar to shapes in training set can be generated by equation (2):

$$x = \overline{x} + Pb , \qquad (2)$$

where $\mathbf{P}$ is a matrix composed of t largest eigenvectors (eigenvectors with largest corresponding eigenvalues) and $\mathbf{b}$ is a vector of model parameters.

The model also should build an average description of image local texture, surrounding every key point. These descriptions of landmark local texture are required for the model to be able to locate landmark positions on new image of modeled object. Determination of such descriptions is very important for further model performance. In [3] it was proposed to use normalized intensity gradient dg_{ij}^n of gray-level profiles along normal to contour tangent:

$$dg_{ij}^n = \frac{dg_{ij}}{\sum_k dg_{ijl}} . \qquad (3)$$

where dg_{ij} is a j^{th} landmark intensity gradient profile of i^{th} shape and dg_{ijl} is a l^{th} component of dg_{ij}. Average description of local texture around j^{th} landmark is computed as following: $d\overline{g}_j = \frac{1}{N} \sum_{\mathrm{i}} dg_{ij}$, where N is a train set size.

Thus, output model consists of mean shape $\overline{x}$, matrix $\mathbf{P}$ and a set of landmark local texture descriptors $\left\{ d\overline{g}_j \right\}$.

Model utilization. Due to possible variations of object image shape, model should be able to tune its parameters to match those shape variations. To match given image of object with obtained model in addition to model parameters $\mathbf{b}$ we should also determine scale s, translation (t_x, t_y) and rotation θ of the shape generated by model (1).

Initial guess about the model parameters $\mathbf{b}$ and position (s, t_x, t_y, θ) should be made for satisfactory final convergence of the model to a given new image (in present paper the initialization of the (s, t_x, t_y, θ) is out of consideration and assumed to be known). After the initial model (i.g. model consisted of mean shape $\overline{x}$, zero vector parameter $\mathbf{b}$ and initial guess for translation, rotation and scale parameters) is placed on new image the search procedure begins. It can be divided on three steps:

1. Search for a new locations x_{new} of key points.
2. Fix b and determine (s, t_x, t_y, θ) that minimize $(x_{new} - x_{cur})W(x_{new} - x_{cur})^T$.
3. Fix (s, t_x, t_y, θ) and find b that minimizes $(x_{new} - x_{cur})W(x_{new} - x_{cur})^T$,

where W is a weighting diagonal matrix (4), which is used to give more significance to those landmark points that have more stability in their displacements with respect to all other landmark points presented in a given shape.

$$W = \begin{vmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & w_n \end{vmatrix}, \tag{4}$$

where w_k is a weight of k^{th} landmark computed as following:

$$w_k = \left(\sum_m V_{km} \right)^{-1}, \tag{5}$$

where V_{km} is a variance of distance between k^{th} and m^{th} landmark points in training set.

In step 1, determination of new locations x_{new} of key points is made by nearby search around current positions of key points. It aims to find better match of average normalized gradient profile $d\overline{g}_j$ (inferred from training images) and normalized gradient profile dg_j^n of input image (computed by (3)) for every landmark j. In this work, search of new locations x_{new} is made by similar procedure to that described in [4].

In step 2, having x_{new} and fixing $\mathbf{b}$ we try to align $x_{current}$ to x_{new} by finding $(s, t_x, t_y \theta)$ that minimizes $(x_{new} - x_{cur})W(x_{new} - x_{cur})^T$. This is done by similar procedure as for training shapes alignment.

In step 3, after $(s, t_x, t_y \theta)$ is determined and fixed we compute new values of model parameters $\mathbf{b}_{new}$ trying to minimize $(x_{new} - x_{cur})W(x_{new} - x_{cur})^T$. Such minimization with respect to model parameters $\mathbf{b}$ is provided by the following:

$$b_{new} = P^{-1}(x_{new} - \overline{x}). \tag{6}$$

The above three steps are repeated until convergence is reached.

Thus, for testing we should provide the model with a new image of modeled object and initial guess for model parameters $\mathbf{b}$ and position (s, t_x, t_y, θ). After model converges, it provides as with a set of output parameters: final model parameters $\mathbf{b}_{final}$ (that describe shape of modeled object) and final position parameters $(s^{final}, t_x^{final}, t_y^{final}, \theta^{final})$ (that describe relative scale, translation and rotation of shape of object depicted on input image and shape generated by (2) with $\mathbf{b}_{final}$ as model parameters).

Active shape model with adaptive weights. Considering ASM applied to radiographic images (that can be characterized as noisy with fuzzy edges of objects) more attention should be paid to building characteristics of image area around every key point. Due to fuzzy nature of radiographic images it can be sometimes difficult to determine the true location of the landmarks. It can cause some landmark point locations to be remote from their true positions on the image. Consequently the whole model can converge in a wrong way, especially when points with big weight values are localized at false positions.

Usage of adaptive weights $\tilde{w}_k$ that allow ASM to avoid problems that could appear with misdetected key points locations, is proposed in this paper. The idea of adaptive weights consist of making weight magnitude dependent on matching rate of average normalized profiles $d\overline{g}_j$ (inferred from training set) and normalized profiles dg_j^n of input image. Such dependence makes it possible to reduce the influence of mislocated points having big weights on final convergence.

Expression (7) is used by conventional ASM to compute the similarity between dg_j^n, that characterize texture of input image around j^{th} landmark point and $d\overline{g}_j$, inferred from training set:

$$f(dg_j^n) = (dg_j^n - d\overline{g}_j)\Sigma_j^{-1}(dg_j^n - d\overline{g}_j)^T , \qquad (7)$$

where $f(dg_j^n)$ is a Mahalanobis distance and Σ_j is a covariance matrix of all normalized gradients dg_{ij}^n that belongs to j^{th} landmark. Thus to make weights adaptable we propose to scale the weights w_k by the $f(dg_j^n)$ (7). Consequently the scaled weights (5) are computed by the following:

$$\tilde{w}_k = w_k f(dg_j^n) . \qquad (8)$$

With adaptive weights $\tilde{w}_k$ model becomes less sensitive to wrongly located during search procedure key points. Even if some landmarks have w_k with large magnitude and are misplaced from its true location, the further scaling by (8) will reduce the magnitude of that weight with respect to all other weights, preventing the model convergence to a false shape.

Experimental results and conclusions. Proposed model performance evaluation and all experimental results were obtained by training and testing it on a set of radiographic images of pipe welds. An example of that type of images is shown at Fig. 1*a*.

For model building a set of 22 manually annotated images was used. After the training, model containing only a small set of significant model parameters (in the given case vector **b** contains only three parameters) were obtained. Despite of that small number of parameters the model is able to reproduce about 99% of training data variation.

After the input image is given, initial model (in our case model with zero parameter vector **b** and guess for scale, rotation and translation parameters provided by user) should be placed on it. An example of initial model placed on input image is shown at Fig. 1*b*. Two model performances were tested when given the same initialization (Fig. 1*b*) of the model. First model was conventional ASM. As depicted at Fig. 1*c* the convergence of the model is not satisfactory for a given initialization. Bad convergence was caused by the misslocation of key points with big weights. In contrast to conventional ASM, proposed adaptive weights ASM converged to plausible final convergence. The result of adaptive weights ASM performance is shown at Fig. 1*d*.

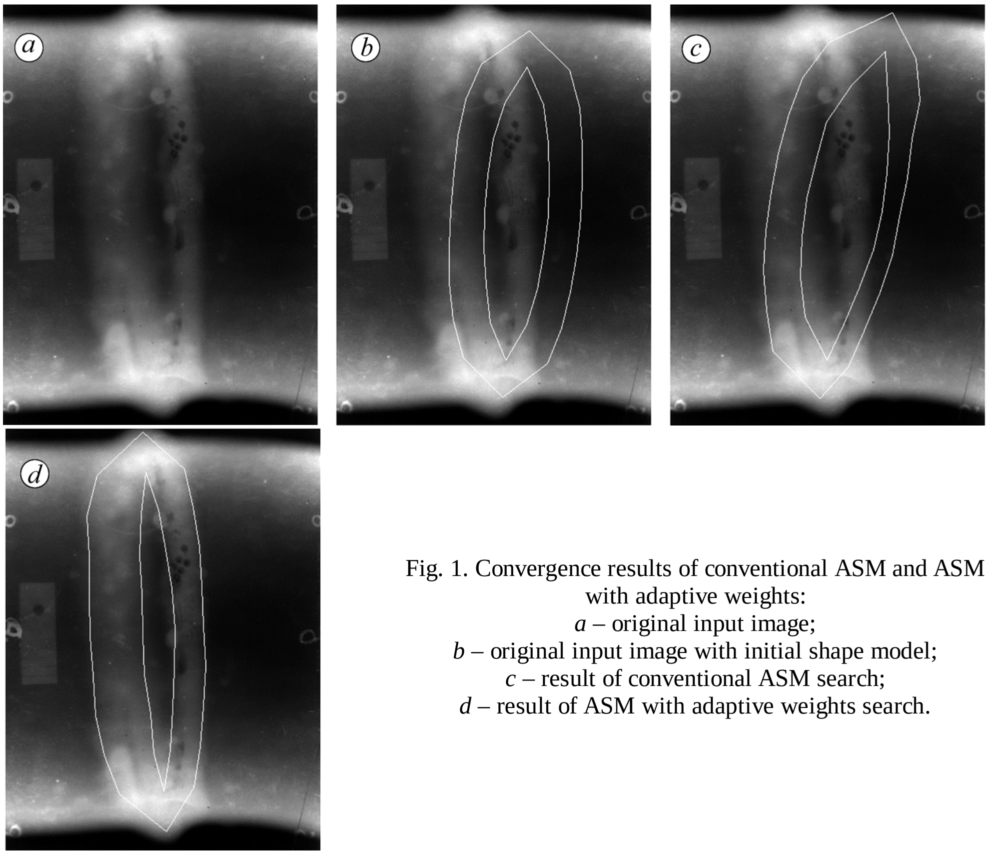

Fig. 1. Convergence results of conventional ASM and ASM with adaptive weights:
a – original input image;
b – original input image with initial shape model;
c – result of conventional ASM search;
d – result of ASM with adaptive weights search.

Experiments show that even though 85...95% of key points are located in proper way the model still can be lead to a wrong convergence by the small amount of key points with a big weights values. This drawback of conventional ASM can be avoided by using adaptive weights values witch are able to change during optimization stage.

This approach also can be used for occlusion handling. For instance, it can be usefull in situations when modeled object tend to appear in cluttered environment where some of its parts are occluded.

Obtained results show the advantage of the adaptive weights ASM utilization over conventional ASM for images with fuzzy edges of modeled objects.

1. *Kass M., Witkin A., and Terzopoulos D.* Snakes: Active contour models // International Journal of Computer. – Vision., 1987. – **1**, № 4. – P. 321–331.
2. *Junmo Kim, Müjdat Çetin, Willsky Alan S.* Nonparametric shape priors for active contour-based image segmentation // Signal Processing. – 2007. – 87(12). – P. 3021–3044.
3. *Region-based* active contours with noise and shape priors / F. Lecellier, S. Jehan-Besson, M. J. Fadili, G. Aubert, M. Revenu, E. Saloux // IEEE ICIP. – USA: Atlanta, 2006. – P. 1649–1652.
4. *Active* shape models – their training and application / T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham // Computer Vision and Image Understanding. – 1995. – 61(1). – P. 38–59.
5. *Double* contour active shape models / M. Seise, S. J. McKenna, I. W. Ricketts, and C. A. Wigderowitz // British Machine Vision Conference. – 2005. – Vol. 2. – P. 159–168.
6. *Learning* Active Shape Models for Bifurcating Contours / M. Seise, Stephen J. McKenna, Ian W. Ricketts, C. A. Wigderowitz // IEEE Trans. Med. Imaging. – 2007. – 26(5). – P. 666–677.
7. *Romdhani S., Gong S., and Psarrou A.* A Multi-View Nonlinear Active Shape Model Using Kernel PCA // Proc. British Machine Vision Conf. – 1999. – P. 483–492.

H. V. Karpenko Physico-Mechanical Institute NAS of Ukraine

Received
24. 05.2008