Multiple model estimation for the detection of curvilinear segments in medical X-ray images using sparse-plus-dense-RANSAC

Please check the document version of this publication:

Multiple model estimation for the detection of curvilinear segments in medical X-ray images using sparse-plus-dense-RANSAC

Chrysi Papalazarou
Univ. of Technol. Eindhoven
Eindhoven, the Netherlands

Peter M. J. Rongen
Philips Healthcare
Best, the Netherlands

Peter H. N. de With
CycloMedia / Univ. Technol.
Eindhoven, the Netherlands

Abstract

In this paper, we build on the RANSAC method to detect multiple instances of objects in an image, where the objects are modeled as curvilinear segments with distinct endpoints. Our approach differs from previously presented work in that it incorporates soft constraints, based on a dense image representation, that guide the estimation process in every step. This enables (1) better correspondence with image content, (2) explicit endpoint detection and (3) a reduction in the number of iterations required for accurate estimation. In the case of curvilinear objects examined in this paper, these constraints are formulated as binary image labels, where the estimation proved to be robust to mislabeling, e.g. in case of intersections. Results for both synthetic and real data from medical X-ray images show the improvement from incorporating soft image-based constraints.

1. Introduction

In this paper, we study model estimation, motivated by the problem of object detection in medical X-ray imaging. Many surgical instruments present in medical images can be described by simple one-dimensional models, for example needles, catheters, guidewires, etc. By making a sparse representation of the image in terms of interest points, we can create the input points for the model fitting, while the original dense data remain available. In many cases, an unknown number of objects can be present in the image, while overlapping objects may make it difficult to assign a model to points at the crossings. Finally, it is not sufficient to only detect the instances of curvilinear objects, but it is clinically relevant to localize their endpoints correctly.

Model fitting in outlier-rich data is an important task in computer vision, especially useful in, e.g., geometry estimation and object detection. A well-known method to perform model estimation in cases of a high outlier ratio is RANSAC [1]. The original RANSAC framework by Fischler and Bolles has seen many modifications and improvements, e.g. [7, 6], aiming at selecting one (the most prominent) model. For multiple model estimation, the simplest approach is to sequentially estimate models and remove their inlier sets from the data [2]. Zuliani *et al.* [8] proposed a parallel multi-RANSAC algorithm that was reported to be more stable and perform better at the correct detection of inliers. However, they used a fixed (known) number of models and assigned each point to exactly one model in a greedy approach, which does not handle the case of intersecting models.

Toldo and Fusiello [5] recently proposed a different framework for multiple model estimation. In their approach, clustering of the hypotheses is used to fit multiple models by assigning a "preference set" to each data point. This preference set expresses the models to which each data point gives consensus. Clustering is performed using the Jaccard distance between all preference sets to select the final models. This method handles intersecting models in a natural way and does not require *a-priori* knowledge of the number of models.

RANSAC was proposed as an improvement to least-squares fitting of a model to sparse data in the presence of outliers. Pre-processing steps, such as interest point detection, generate the sparse data necessary to perform the fitting. This pre-processing typically involves a loss of information that influences the quality of the fitted model. We observe that both the quality and the efficiency of the fitting can be improved by exploiting the available dense image data.

Given the unknown number of objects and the sparse distribution of interest points, the object detection requires a flexible and robust multi-model estimation, which preferably exploits all available information. We employ a RANSAC-based scheme that uses both sparse and dense information, and term it sparse-plus-dense(SPD)-RANSAC. The model generation step is guided by a rough segmentation of the image into la-

IEEE
computer society

bels representing possible locations of elongated objects, which speeds up the selection of good candidate models. To allow for multiple model assignments for points on crossing segments, a double run of consensus query is applied. The labels are also used to impose correct endpoints on the segments.

Section 2 discusses the preprocessing to obtain the sparse plus dense input. In Section 3 the multiple model estimation method is described. Experimental results are given on synthetic and real X-ray data in Section 4, while a discussion of the results follows in Section 5.

## 2. Generation of sparse plus dense data

Sparse data: The considered images in our study are medical X-ray images of curvilinear objects, e.g. needles or catheters, overlaid on the anatomy. The preprocessing starts with a bothat filter to smooth out background variations. Afterwards, a Gaussian scale-space is calculated and a ridgeness representation R of the image is formed, as in [3]. The feature points are then selected as maxima in the scale-space of R, and form the sparse input to the model fitting.

Dense data: An additional output of this step is a label image, indicating prominent elongated objects in the image. This is formed in the following way. First, an image of the rms second derivatives at the largest scale $I_{D2} = \sqrt{Lxx^2 + Lyy^2}$ is created. This image is morphologically reconstructed, using $0.5I_{D2}$ as the marker, and thresholded using adaptive thresholding. Connected-component labeling is then applied to this binary image to obtain candidate curvilinear structures. It is important to note that this label image does not need to capture the topology entirely accurately, as will be shown in the next section. The labels are rather used as a *soft* constraint to guide the model fitting and assist the correct endpoint detection. An example of these steps can be seen in Figure 1.

3. Multiple model estimation

In the model estimation step, we can discriminate between hypothesis generation and consensus query. Additionally, we apply a second run for the consensus query to allow multiple model assignments per data point. The analysis is performed for the case of line-segment estimation, but can be generalized to other 1D curves such as polynomials or splines (see e.g. [4]), by changing the estimation and cost functions. In the following, names in italics refer to function names of Algorithm 1.

Hypothesis generation: The first sample is chosen randomly among the input points that are lying on a
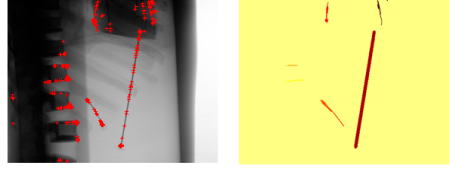
Figure 1. Example of input generation. Left: original image with ridgeness maxima. Right: label image (color coded).

non-zero label. For the remaining samples, the probability of being selected varies with the distance of that point to the skeleton of the label of the first point:

$$P(\mathbf{x}_j \| \mathbf{x}_i) = \frac{1}{Z} \exp\left(-D(\mathbf{x}_j, skel(L(\mathbf{x}_i)))/\sigma_L\right), \quad (1)$$

where $D(\mathbf{x}_j, L(\mathbf{x}_i))$ denotes the (Euclidean) distance of $\mathbf{x}_j$ from the skeleton $skel$ of the label $L(\mathbf{x}_i)$ where $\mathbf{x}_i$ lies, obtained through morphological skeletonization of $L(\mathbf{x}_i)$, Z is a normalization constant, and σ_L expresses the attraction range of the labels.

In the proximity criterion of [2, 8], the assumption is made that the average *inlier-inlier* distance is smaller than the average *inlier-outlier* distance and thus, sampling probabilities should be weighted by a proximity distribution. In our sampling scheme, the sampling probability is weighted by the distances of the points to each *label* (procedure *GetMSS*). The assumption here is that points with the same label have a higher probability of belonging to the same model, so hypotheses should be formed more frequently among these points.

The total number of iterations needed to form hypotheses for all models is equal to the sum of iterations needed for each model k, so that $M_{tot} = \sum M_k$. For simplicity of the following analysis, we assume a uniform sampling distribution among all points for a given label[1]. Let N_l be the total number of points on label $l = 1..L$, S_{min} the smallest number of inliers and d the cardinality of the minimal sample set (MSS), the minimum number of points required to estimate a model (e.g. for lines it is 2). As in [5], the probability of producing an outlier-free MSS from N_l in i samplings is:

$$P(E_i \| E_1, E_2, ..., E_{i-1}) = \frac{S_{min}!(N_l - i)!}{(S_{min} - i)!N_l!}. \quad (2)$$

Then, $p = \prod_{i=1}^{d} P(E_i \| E_1, E_2, ..., E_{i-1})$ can be approximated as:

$$p \approx \prod_k \frac{(S_{min} - d + 1)e^{-\alpha_L^2/\sigma_L}}{(N_l - S_{min} - d + 1)e^{-\omega_L^2/\sigma_L} + (Smin - d + 1)e^{-\alpha_L^2/\sigma_L}}, \quad (3)$$

---

[1]In fact, the sampling probability is weighted with the ridgeness of each point to encourage earlier selection of stronger feature points.

```
input    : X (sparse data), L (dense data), Q (iteration
           threshold), R (probability weight)
output   : Multiple models theta, explaining X, and
           their respective consensus sets CS

X_left = X; k = 1;
while X_left is not empty do
    theta(k) = GetMSS(Xleft, L);
    CS_0 = GetCS(theta(k), X_left);
    CS_1 = GetSegmentPoints(CS_0, L)
    X_left = RemoveCSFromInput(X_left, CS_1);
    CS_x = ExtendCS(theta(k), X);
    CS(k) = GetSegmentPoints(CS_x, L)
    for l ← 1 to k do
        theta(l), theta(k), k =
        CheckMerge(theta(l),
        theta(k), CS(l), CS(k));
    end
    if X_left has not changed in last Q iterations then
        break;
    end
    k = +1;
end
Procedure [theta, MSS] = GetMSS(X, L, R)
points_on_label = OnLabel(X, L); // Returns
    points with non-zero label
sample1 = SampleWithProb(points_on_label, R);
    // Samples with a given probability R
L_i = L(sample1); // label on which sample1 lies
D = DistSkel(L_i);// Distance transform from
    skeleton of L_i
P = DistanceProbability(D); // Equation 1
next_samples = SampleWithProb(X − sample1, R);
MSS = Concatenate(sample1, next_samples);
theta = EstimateModel(MSS)
```

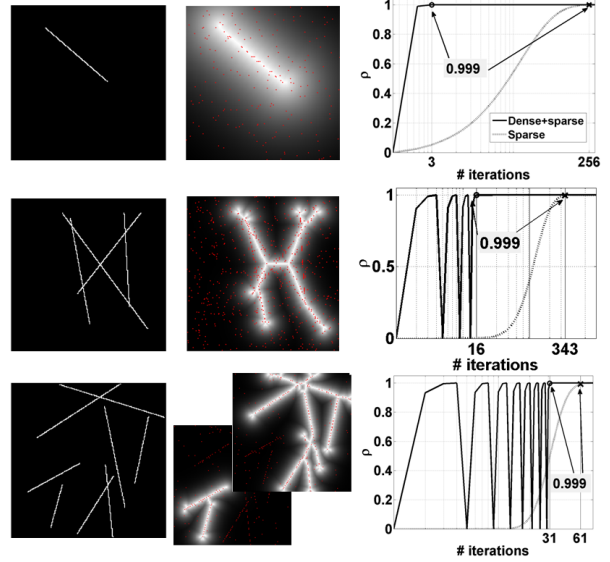**Algorithm 1**: The proposed multiple-model esti-
mation algorithm

Figure 2. Synthetic example for $K = 1, 4, 8$.
**Left: original image. Middle: Result of
Eq. 1 with overlaid input points (for** $K = 8$
two labels are shown). Right: ρ **vs** M_{tot}
**using Eq. 1, compared to pairwise dis-
tances. X-axis in log scale.**

where α_L, ω_L is the average distance of an inlier (re-
spectively outlier) to the label of the correct (respec-
tively false) model, and we set $\sigma_L = 5 \max(sigma_D)$,
the largest scale of the scale-space. In M_k iterations of
the sampling, the probability of obtaining J outlier-free
MSSs is:

$$\rho(M_k) = 1 - \sum_{k=0}^{J-1} \binom{M_k}{k} p^k (1-p)^{M_k-k}. \quad (4)$$

In Figure 2, a synthetic example is given to show
the effect of this sampling scheme, compared to using
pairwise proximity. For the latter, $\sigma = \sqrt{\alpha^2 + \omega^2/6}$
was used, similar to the experiment of [5][2]. For our
sampling scheme, Eq. 4 was run K times with $J = 1$

----

[2]In their experiment, α and ω were functions of σ; here they are
computed from the synthetic data, so that σ was chosen between the
values of [5].

and M_k was chosen as $arg(\rho(M_k) = 0.999)$, to ob-
tain M_{tot}. The gain in the number of iterations using
dense information is apparent. This gain decreases as
the number of models increases. This is because the
label images, as defined in Section 2, may fail to give
unique labels to each model (and thus provide appropri-
ate weights to the correct points) in cases of many in-
tersecting models. This is an issue of the preprocessing
stage, which can be improved by an appropriate choice
of labeling function. However, even for $K = 8$ there is
a speed-up by a factor of two in our experiment, while
for $K = 1$ the gain is two orders of magnitude.

Consensus query: The first consensus set CS_0 is
formed by all points with Euclidean distance below
the RANSAC threshold. To impose correct endpoints,
only points lying inside the corresponding label are kept
(procedure *GetSegmentPoints*). These points (CS_1) are
then removed from the input for the next iteration, so
$\{X_{left}\} = \{X\} - \{CS_1\}$ (see procedure *RemoveCon-
sensusFromInput*).

**Extension of the consensus set:** A second consen-
sus query is then performed in procedure *ExtendCon-
sensus*, using all original input points X to get CS_x.
This allows points to be assigned more than one model.
Again, points outside the labels are rejected and the fi-
nal consensus set for this model, CS, is obtained. The
process is repeated for the remaining input points X_{left}
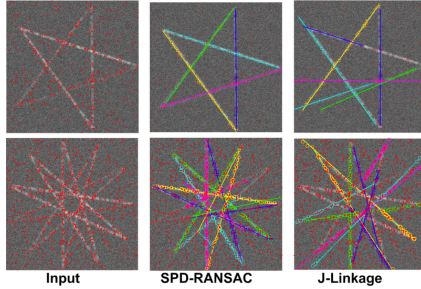
Figure 3. Results on *star* data. Top: 2000 iterations, Bottom: 3300 iterations

to estimate the next model, until no more new models are found for at least Q iterations. For our data, $Q = 3$ was sufficient throughout our experiments.

Model merging: In case two detected models are very similar, they may be referring to the same instance. For each new model k created, possible merging with a previous model l is checked. If the average error of the consensus set of l, CS^l given model k, as well as the average error of CS^k given model l, are below the RANSAC error threshold, then a new model is estimated using the union of CS^k and CS^l.
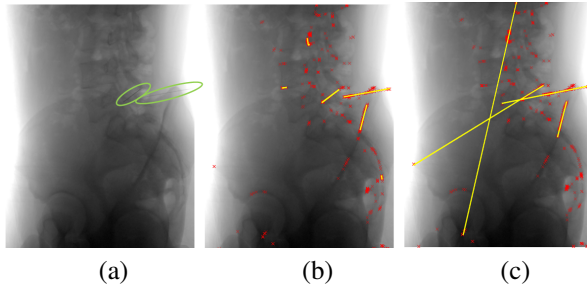
Figure 4. Clinical image results: (a) Original, needles highlighted, (b) SPD-RANSAC, (c) J-Linkage (both 1500 iter's).

4   Experimental results

The synthetic *star* dataset of [5] was used to create dense data, by fitting line segments of different contrasts $(9 - 100\%)$, blurring and adding both signal-dependent and signal-independent noise to simulate the combination of photon and electronic noise present in X-ray images. The corruption of the image affects the quality of the label image. To show the effect of using dense information, we have compared (Figure 3) with the implementation of [5], where only sparse points are used. For the same number of iterations, it is seen that both detection and localization of the models is better for SPD-RANSAC, even in the case of $K = 11$. Merging of similar models reduces the number of false positives, while the localization of endpoints is explicit.

Results on a clinical dataset from a vertebroplasty procedure are shown in Figure 4. The clinical image contains two surgical needles of low contrast-to-noise ratio, which are difficult to distinguish from the neighboring anatomy. The soft constraints ensure better correspondence with image content, while not allowing false models to arise that are irrelevant to the image structure. Moreover, this result demonstrates a case were the correct endpoint detection achieved by the proposed method is crucial.

5   Conclusions

We have presented a method to combine dense and sparse image data for model fitting. Our approach can be regarded as a fusion between segmentation and model estimation, where the results of a rudimentary image segmentation are used to guide the estimation. An implementation of the approach was shown for the problem of curve fitting in medical images, where the high contrast-to-noise ratio and the complexity of the content require robust model fitting. The RANSAC principle is used for the model estimation, where each step is guided by a dense feature representation. We have shown that this leads to faster convergence to plausible models and better correspondence with the image structure. Additionally, it facilitates the correct detection of endpoints, which is a clinically important task.

References

[1] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM*, 24(6):381–395, 1981.

[2] Y. Kanazawa and H. Kawakami. Detection of planar regions with uncalibrated stereo using distributions of feature points. In *BMVC*, pages 247–256, 2004.

[3] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever. Evaluation of ridge seeking operators for multimodality medical image matching. *IEEE Trans. PAMI*, 18(4):353–365, 1996.

[4] C. Papalazarou, P. Rongen, and P. H. N. de With. Surgical needle reconstruction using small-angle multi-view X-ray. *Submitted to ICIP 2010*.

[5] R. Toldo and A. Fusiello. Robust multiple structures estimation with J-Linkage. In *ECCV (1)*, pages 537–547, 2008.

[6] B. J. Tordoff and D. W. Murray. Guided-MLESAC: Faster image transform estimation by using matching priors. *IEEE Trans. PAMI*, 27(10):1523–1535, 2005.

[7] P. H. S. Torr and A. Zisserman. MLESAC: a new robust estimator with application to estimating image geometry. *Comput. Vis. Image Underst.*, 78(1):138–156, 2000.

[8] M. Zuliani, C. S. Kenney, and B. S. Manjunath. The multiRANSAC algorithm and its application to detect planar homographies. In *ICIP*, Sep 2005.