OPTIMAL PARAMETER SELECTION FOR MEANSHIFT TYPE SEGMENTATION

Dragos Bozdog[1], Ionut Florescu[1], Rustam Stolkin[2]
[1]Department of Mathematical Sciences, Stevens Institute of Technology, USA
[2]School of Computer Science, University of Birmingham, UK

LOGO

MOTIVATION

Many vision papers are concerned with development of the best and the fastest computer vision algorithm. However, in many cases when independently implemented these algorithms fail to perform as advertized. The reason is that, in the vast majority of situations vision algorithms are very sensitive to the choice of parameter values used. Chose the parameters detailed in the paper but apply the algorithm to a different image sequence than the one presented in the article and very likely you have an underperforming algorithm.

PURPOSE:
Study the possibility of adapting an algorithm's parameters with respect to the current image specific information.
For convenience we chose to study the Click! algorithm presented on the parallel slide.

SPECIFIC OBJECTIVES OF THE STUDY

The objectives of the statistical analysis:

1. Determine if there exist parameters that can be studied separately from the other parameters.
2. Determine if there exist a set of parameters which are optimal for all the objects within a image
3. Determine if there exist a set of parameters optimal for all the objects in all the images.
4. Determine relations between the input conditions and the optimal parameter choice

THE PARAMETERS

Study the possibility of adapting an algorithm's parameters with respect to the current image specific information.
For convenience we chose to study the Click! algorithm presented on the parallel slide.

We are studying the influence of four parameters, two originally present in the meanshift algorithm:

h_r = The range of pixels in the image included in the meanshift calculation

h_c = The range of pixels in the color space included in the meanshift calculation,

and two new ones introduced by our specific algorithm

Initbox= Size of the initial selection containing object pixels

α = Confidence level defining the ellipse level set

We consider the following levels for each parameter:

- *Initbox*: {3,5,7,9,11}
- h_r: {3,5,7,9,11}
- h_c: {1,2,3,4,5,10,20,50}
- α: {95%, 97%, 99%}

The number of possible combinations is 600. We repeat the segmentation for 10 different randomly chosen starting points within each object, which gives a total of 6,000 data points for each object. We analyze a total of 23 objects in 5 images and thus the total number of observations in our study is 138,000. For each object chosen, the Berkley data set records the true segment, as determined by human operators. Thus, for each of our data points we run the segmentation algorithm and we record two types of errors:

Error I = number of object pixels erroneously classified by the algorithm as background
Error II = number of background pixels erroneously classified by the algorithm as object

Then we calculate a response variable Y as the total error expressed as a proportion of total object size:

$$Y = \frac{Error\ I + Error\ II}{ObjectSize}$$

Clearly, the two types of error are fundamentally different and could have penalized more one type or the other, however for the current analysis we decided to penalize them equally.

THE IMAGES UNDER STUDY

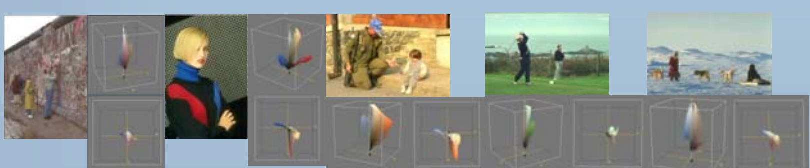

Figure 1. Image 1 and the original LUV space representation. Situation: small objects in cluttered background. The five selected objects are: 1- blue pants for man, 2-blouse for man, 3 – yellow trench for older woman, 4- brown coat and 5 - skirt for faraway woman. The LUV space is very elongated and compact.

Figure 2. Image 2. Situation: Large objects homogeneous in color. No distraction from background. The five selected objects are: 1- right red patch on pullover, 2-blonde hair, 3- blue pullover neck, 4-hand, and 5- left blue patch on the pullover. The LUV space is very wide and non-compact.

Figure 3. Image 3. Differently sized objects on cluttered background. The four selected objects are 1-child one piece suit, 2 - child head, 3- soldier hat and 4 - soldier uniform. The LUV space is relatively compact but wide.

Figure 4. Image 4. Similarly sized objects on distinct background. The four selected objects are: 1- blouse and 2 - pants for the viewer in the back; 3- blouse and 4 - pants for the golf player. The LUV space is very compact and elongated and the object colors represent a small fraction of the total space.

Figure 5. Image 5. Small objects on clear background. The five selected objects are: 1- leftmost fox colored dog, 2- Eskimo in red dress, 3- brown dog, 4- the two white dogs, 5- Eskimo in black dress. The LUV space is wide and relatively non-compact.

STATISTICAL ANALYSIS

Objective 1:
We perform 4 way ANOVA with interaction to determine if we can simplify analysis
X_1=*Initbox*, $X_2=h_c$, $X_3=h_r$, $X_4=\alpha$

Img	Obj	X1	X2	X3	X4	X1*X2	X1*X3	X1*X4	X2*X3	X2*X4	X3*X4	X1*X2*X3	X1*X2*X4	X1*X3*X4	X2*X3*X4	X1*X2*X3*X4
1	1	X	X	X	X	X	X	X	-	X	X	-	-	-	-	-
1	2	X	X	X	X	X	-	X	X	X	X	-	-	-	-	-
1	3	X	X	X	-	X	X	X	-	X	-	-	-	-	-	-
1	4	X	X	X	X	X	X	X	-	X	-	-	-	-	-	-
1	5	X	X	X	X	X	X	X	X	X	X	-	X	-	-	-
2	1	X	X	X	X	-	X	X	X	X	X	-	-	-	-	-
2	2	X	X	X	X	-	X	-	-	X	X	-	-	-	-	-
2	3	X	X	X	X	-	X	-	X	-	-	-	-	-	-	-
2	4	X	X	X	X	X	-	-	X	X	-	-	-	-	-	-
2	5	X	X	X	X	X	X	X	-	X	X	-	-	-	-	-
3	1	X	X	X	X	X	X	-	X	X	X	-	-	X	X	-
3	2	X	X	X	X	-	X	X	X	X	X	-	-	-	-	-
3	3	X	X	X	X	X	X	X	X	X	X	-	X	-	X	-
3	4	X	X	X	X	X	-	X	X	X	X	-	-	-	-	-
4	1	X	-	X	X	-	-	X	-	-	-	-	-	-	-	-
4	2	X	X	X	X	X	X	X	X	-	-	-	-	-	-	-
4	3	X	X	X	X	X	X	X	X	-	X	-	-	-	-	-
4	4	X	-	-	X	X	X	X	-	-	-	-	-	-	-	-
5	1	X	X	X	X	X	X	X	X	X	X	X	X	X	-	-
5	2	X	X	-	X	X	X	X	-	X	-	-	-	-	-	-
5	3	X	X	X	X	X	X	X	X	X	-	-	X	-	-	-
5	4	X	X	X	X	X	X	X	X	X	X	X	X	X	-	-
5	5	X	X	X	-	X	-	X	-	-	-	-	-	-	-	-

Answer:
We cannot eliminate interaction terms (checkmark denotes term is statistically significant).

All 4 parameters have to be studied together

Objective 2:
We did find parameters that work for all objects in the image, with the exception of image 3. Here are the optimal values:

	$(h_r, Initbox, h_c, \alpha)$	$(h_r, Initbox, h_c, \alpha)$	$(h_r, Initbox, h_c, \alpha)$	$(h_r, Initbox, h_c, \alpha)$
Image 1	(9, 11, 3, 95%)			
Image 2	(3, 7, 10, 99%) (3, 7, 20, 99%) (3, 7, 50, 95%) (3, 7, 10, 97%)	(5, 7, 1, 99%) (5, 7, 10, 99%) (5, 7, 20, 99%) (5, 7, 50, 97%)	(9, 7, 50, 97%)	(11, 9, 20, 95%) (11, 9, 20, 97%)
Image 3	none			
Image 4	(3, 3, 50, 99%)	(3, 5, 20, 99%) (3, 5, 50, 97%)	(5, 3, 50, 99%)	
Image 5	(3, 3, 10, 97%)	(9, 3, 10, 99%)	(9, 5, 5, 95%) (9, 5, 5, 97%) (9, 5, 5, 99%)	(11, 3, 10, 99%) (11, 7, 4, 95%) (11, 7, 4, 97%)

Objective 3:
It is not possible to use the same parameters working well in all situations. Furthermore, looking at image 3 we see that even in the same image the optimal set is varies from object to object.

Objective 4:

This preliminary study indicates that relations between local measurements around the segmented object are beneficial and could be put in relation with the optimal parameter choice. In particular we found that the following measurements would help choose the best parameters:

- A local measure of clutter
- A local measure of variation of color histogram
- A degree of texture change

Situation	h_r, *Initbox*	h_c	α
Small objects, homogeneous color, cluttered background	large	small	small
Large objects, homogeneous color, clear background	small	large	large
Small objects, non-homogeneous color, clear background	large	large	any
Small objects, homogeneous color, clear background	small	small	Small