Optimizing color information processing inside an SVM network

J. Pasquet $^{1,3}$, G. Subsol $^3$, M. Derras $^1$ and M. Chaumont $^{2,3}$

$^1$Berger-Levrault, Labège, France
$^2$University Nîmes, France
$^3$LIRMM, University Montpellier / CNRS, France

Visual Information Processing and Communication VII
Electronic Imaging 2016
Outline

1. Motivations
   - Problem
   - Used features
   - The SVM Network

2. How taking into account color information
   - Color space
   - Design proposal for SVM Network

3. Experimental results
Urban Object Detection In Aerial Images

Urban object definition:

- The appearance of the objects varies: color, size, orientation...
- Multiple distortions and occlusions due to shadows, vegetation..
- Multiple-object detection

About the databases of aerial images:

- 19 images for the training database
- 3 images for the validation database
- 2 images for the testing database
Features Extraction

Extraction of multiple Histograms of Oriented Gradients

- Calculate HOG within cells and blocks
- Accumulate features to construct HOG descriptor of different sizes

Figure: HOG representation for 9x9 cells size.
How can we extract the best descriptor based on the HOG descriptor when there are 3 channels (color image)?

1. Transform the color image input into grey level
2. Extract HOG descriptor on each channel and concatenate them
3. Only take into account the highest gradient
4. Reduce the dimension (PCA...)

Color Features Extraction
HOG and SVM Network

Figure: SVM Network works on grey level images.
HOG and SVM Network

Figure: SVM Network works on color images called separate architecture.
Color space

What is the best discriminant color space?

- RGB
- $Y = 0.21R + 0.71G + 0.07B$
- CIE — LAB and CIE — LUV (based on human perception)
- HSV (cylindrical-coordinate representations of RGB)
- Use a PCA transformation
Experimental results

SVM Network vs SVM Single:

- SVM Network always improves the precision
- The performance from different color spaces are very closed.
- HSV and RGB: the 2 bests

=> But the colors are threatened separately in the SVM Network...
Experimental results

SVM Network vs SVM Single:
- SVM Network always improves the precision
- The performance from different color spaces are very closed.
- HSV and RGB: the 2 bests

=> But the colors are threaten separately in the SVM Network...
Design the network

1- Fusion of channels per resolution

Concatenate the different channels from the same resolution.

Different channels are not necessarily in the same feature space, e.g. : HSV.

=> Normalisation problem
Design the network

2- A specific function to merge the channels per resolution

Connect all input neurons with the same resolution $r$ to a specific function:

- No focus on the low variations (Maximum function)
- Linear quantification (Product function)
- The first principal component (PCA transformation)

Figure: Representation of the Maximum Network

$\Rightarrow$ The quantification may lose important information.
Design the network

3- Stacking of channels per resolution

- Features from different channels are scaled and used in the same neuron
- Each weight is learned by using an SVM

Figure: Representation of the Stack Network
Figure: ROC curves for the separate and fusion architectures
Results

Figure: ROC curves for the fusion, product and maximum architectures
Results

Figure: ROC curves for the fusion, maximum and stack architectures
Overview of the Network performance

Figure: Precision gain is compared to the separate architecture.
Conclusions

- SVM Network outperforms SVM by an average precision gain ranging from 1.5% to 6%.
- SVM Network design (e.g. Stack Design) would improve to 10% the precision.

Future works

- Design an SVM Network to combine several feature types: SURF, SIFT...
- Design a CNN input neurons.
Thanks!
The best Parameters for SVM Network

Figure: Tuning of the SVM Network on a validation database

Before each experiment the SVM network used a validation database to fix the best number of hidden and random neurons.