MALIS PROJECT

POSE ESTIMATION OF SNEAKERS IN A PICTURE

Eugène Berta, Patrick Iversenc

We intend to work on the following problem :
From a picture containing a sneaker, we want to extract the 3D positioning of the shoe.
We want to obtain the orientation of the shoe in the three dimensional repository induced by the camera. In an ideal scenario, we will be able to create a 3D model of the shoe in a camera based repository.
If we succeed to develop such an algorithm, it can be combined with imaging deep learning algorithms to replace a shoe in a picture or video with any other shoe. This kind of imaging manipulations have many potential commercial uses in the fashion and advertising industries, to create personalized video advertising for example.

We intend to base our work on existing deep learning algorithms which will able us to detect the presence and positioning of a shoe in a given picture.
Based upon this first result, we want to use our knowledge of the shape of the shoe to deduce the pose of the sneaker in space. In order to do this, we consider using different methods :
Our first idea is to use the result of the first algorithm to approximate the shoe's shape in the picture with an ellipse. Using the orientation of the ellipse's main axis and the distance between the two focuses we will deduce the orientation of the shoe in the three dimensional space. This can be done by using linear regression techniques that we studied in the MALIS courses. We will have to find the best way to approximate the shoe with the ellipse, and train multiples algorithms to optimize our results.
We also red about convolutional neural networks algorithms that connects 3D models of objects to pictures of this same object. Using this idea, we may be able to develop a more complex algorithm using a generic 3D model of a sneaker. Our idea is to better approximate the shape of the shoe in the picture using views from different angles provided by the 3D model.

We intend to test our algorithms in different ways :
- Our first deep learning algorithm can be trained and tested using existing datasets accessible online, we will also test it with the pictures that will be used to train our second algorithm in order to get the best level of performance before we start trying new things.
- We will train and test our second algorithm using pictures that we will take ourselves, in a controlled environment where we can measure the camera-sneaker distance and the sneaker's orientation.

Prior research and datasets :

K. Yamaguchi, et al., “Parsing clothing in fashion photographs,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2012, pp. 3570–3577.
Deep fashion database : http://mmlab.ie.cuhk.edu.hk/projects/DeepFashion.html