# New Perception of Fluffy Surfaces

Alexander Kadyrov, Hui Yu and Honghai Liu
School of Creative Technologies, University of Portsmouth, UK

*Abstract*—Scene material recognition is a valuable perceptual ability for robots. However, current methods fail to provide robust solution for indoor surrounding material recognition. As human beings are able to immediately understand properties of surrounding materials, especially when it concerns smooth fluffy materials, we believe robots should be equipped with similar abilities. In this paper we explore a new idea for fluffy surface perception for robots based on above hypothesis. The aim is to enable robots to distinguish smooth and fluffy surfaces without touching them. This is achieved through calculating image matchability map from video cameras. Through measuring the similarity of images captured from different viewpoints, robots are able to immediately recognize whether it is a fluffy material. The method has been validated by primary experiments. Our results show that robots can have a sense of material properties for fluffy materials without touching them or any prior knowledge.

*Index Terms*—Perception, fabric, fluffy material, computer vision.

## I. Introduction

A robot in human environment confronts a great challenge. The robot should look around with human eyes, that is to share (or understand) the elementary basic values with humans, this is the essence of Anthropomatics, that highlights the paramount role of Robotics and Automation in improving human quality of life. The most basic values for us are in materials, because we all surrounded by materials, starting form our clothes on our bodies and the floor where we have to stand on. For humans it is important whether towels are fluffy, lofty and warm. Can a robot visually distinguish them from stiff and scratchy old towels? Such a secondary question becomes important when robots enter our life. People of all ages are rather interested in such abilities of robots. Images of the towels can be almost the same in both cases, then, obviously, more information from video stream is required to provide sufficient information for detecting the properties of such surfaces.

In everyday life humans are surrounded by natural and artificial materials with various textures. New materials and textures enter the human world daily. We can easily classify materials in way that we primarily need, that is to distinguish hard vs. soft, pleasant vs. unpleasant to touch. smooth vs. fluffy surfaces. It is not obvious how we can do this without touching, but it is evident that while assessing a new, not seen before piece of material, people prefer look at it at slightly different angle. Then a decision whether the material is fluffy, comes intuitively.

On the other hand, in computers the decision making process goes by comparison with a given database, and then a computer cannot assess not seen before material. Computers (more precisely, modern algorithms), recognise a *photograph* of the material rather than the material itself. At first glance, it seems that the latter sentence has no much meaning since a mounted camera produces photographs anyway. To clarify the issue, let us remind the case when Android Face Lock feature spoofed by photograph[1], then it made evident that recognition of a *photograph* of a face and recognition of the face are still the same for modern recognition systems, but in reality they are not equivalent[2].

This makes important the observation that a human examines materials "making images" from different angles. This is true that the camera produce a picture, but is also can produce a stream of pictures, the same as our eyes do. This enables us to propose a method to distinguish flat surfaces from fluffy ones using images taken from different angles. Typical fabric has highly uneven surface, and this can be detected by the proposed method. While it would be extremely difficult to reconstruct tiny variations in a surface, it is still possible to assess its fluffy structure integrally. For this, a robot may only have a look at the material from two different locations; similar to how humans do it using two eyes, and then consider whether the two pictures are true warped versions of one another. A smooth surface in the two images matches with itself perfectly, whereas fabric from different views produces less matchable images, and thus the robot gets a clue for fabric segmentation. This newly discovered effect was explained and demonstrated in the paper.

The importance of the proposed method is in need of robots to work with fabric and other fluffy (e.g. fur) materials, when robots have to delineate these materials from smooth surfaces which may have exactly the same texture. Humans solve the latter effortlessly, making difficult even to apprehend this problem for a robot.

In this paper we announce a new approach, show that it works on examples. Currently, the method does pairwise comparison of surfaces in a way as humans do. It would be possible to provide numerical assessment of surfaces after a large experimental work, but it this is not the aim of this paper.

## II. Related work and the suggested approach

Material classification was an active research topic for quite a long period ( [4], [6]) and attracted researchers from different

---

[1]http://www.clubic.com/technologies-d-avenir/actualite-507492-sensiblevision-futur-reconnaissance-faciale.html

[2]Our presented method can solve this particular dilemma; however, here we focus on fluffiness of materials only.

communities including computer vision, machine learning and sensing.

Various techniques have been proposed to address the material classification problem; nevertheless, efforts in this area in computer vision fall into the category of texture classification, which can be regarded as a sub-problem of material classification, whereas the nature of material classification is in fact more complicated. For instance, a classifying system that is designed following the approach of texture classification will classify a printed photograph of a fabric surface and the real fabric surface, to be the same. In other words, material classification remains an open problem and rather a promising topic.

More recently, within the community of computer vision, research on 3D reconstruction emerged and soon became a hot topic. It also suggests some new possibilities to address the material classification problem, ideally by virtue of measuring distance of those object captured by the camera at each pixel. However, current 3D reconstruction techniques do not pose the task to reconstruct tiny variations in a surface. The very basic approach in this direction is just using two or more images taken from different locations. And, indeed, such an approach exists. A broad ideas to examine texture from different angles were tried, among other papers, in [2] and [7]. The general design was to illuminate and view a surface under all possible angles, and analyze each received image, then store obtained histograms for future comparison. However, in [9] it was established that an approach based on only *one* image still can supersede. The idea to use one image only prevails; the trends are analyzed in [6].

We may envisage that an effective usage of 3D information, or, more generally, multiple images of the same patch, remains to be an opportunity in material characterization. On the other hand, a clear need in fabric recognition appeared [8] mainly due to advances in robotics. Thus, the ability to distinguish fabric from other surfaces is required when both kind of surfaces are present in a view of a robot.

In this paper we propose another way to extract space-related information from images. It is neither a reconstructed depth, which is known as 2.5D information, nor a plain 2D image. Rather, in some sense it can be considered in between. The information, which we present it as a *matchability map*, reflects small irregularities within the surface of the material. The idea behind the proposed approach is as follows. Two images of a fluffy surface will contain different information and then cannot precisely coincide when being properly aligned. If even the angle between the views were small, we would see a changed surface. This is because a fluffy surface is irregular and covered with many tiny hairs. Let us illustrate this with a drawing in Fig 2. A profile of an imaginary fluffy surface is depicted and presented by hairs and it is viewed from two cameras $A$ and $B$ places above. One can notice that camera $A$ can see a point that is not seen from the position of camera $B$ (shown by two arrows). So then in this case the images seen by the two cameras cannot be precisely warped to each other; it is impossible, because one image has information that is not
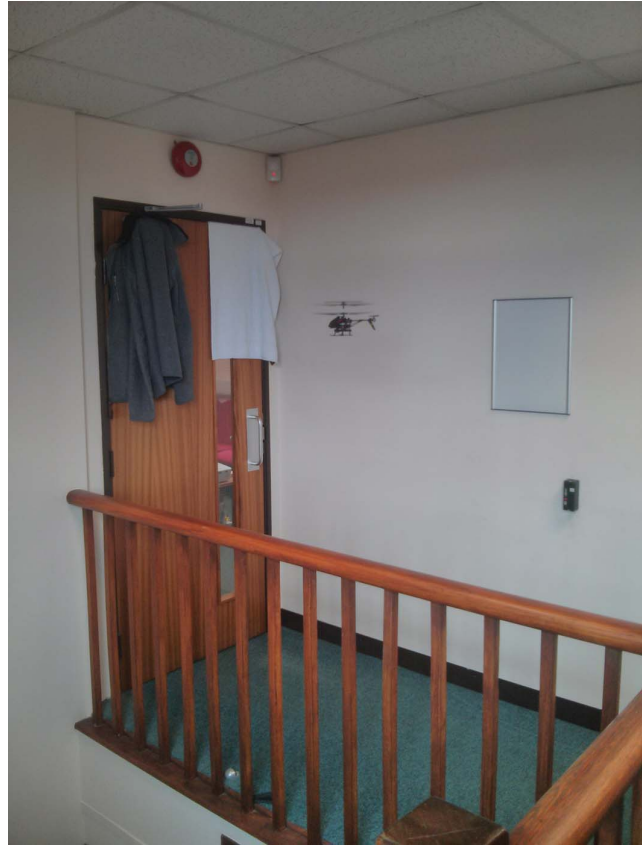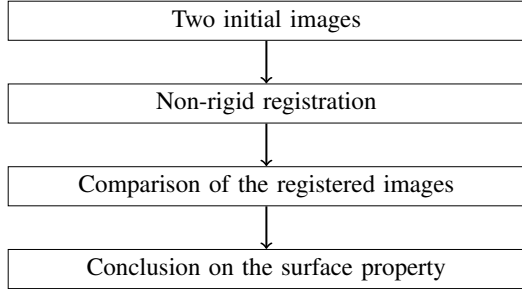


Fig. 1.  Flying robot (a helicopter with a camera, in the center of the picture) takes images while searching fabric as fluffy surfaces.



Fig. 2.  Two views

available in the other image, and vice versa. Nevertheless, the hairs are thin and semi-transparent, then the images are very similar, and this enables us to find a correspondence between the images.

Based on this observation, we may imagine a method for comparison of surfaces in their fluffiness. The idea is this: the two images are registered by an accurate method, and quality of such registration reflects fluffiness. In short, we suggest that images of a patch of a flat surface are more correlated

to each other than images of a patch which belong to a fluffy surface. Note in this paper we use terms "fluffy", "matchable", "fabric", "'smooth" as ordinary English words. For example, fluffy means "covered with very soft hair or feathers". So then, the novelty is the next simple flowchart

```
┌─────────────────────────────────────┐
│         Two initial images           │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│        Non-rigid registration        │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│   Comparison of the registered images │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│   Conclusion on the surface property  │
└─────────────────────────────────────┘
```

This paper aims to present this discovery, explain the effect, and present experiments demonstrating the idea. The rest of the paper is organized as follows. In the next section we formulate the method, describe its essential element which is a registration algorithm. Then we demonstrate the method in two experiments. In the discussion section, we explain possible limitation and other particularities of the method and consider further work. We conclude the paper with a short overview of its results.

## III. METHOD

### A. Matchabilty map

To explain the idea, let us remind that 3D reconstruction starts from finding correspondence between two images of the same scene; ideally, for each pixel in the first image $f$ we should find a corresponding pixel in the second image $g$, and the both pixels depict the same point of the scene. After that, 3D reconstruction can proceed. However, we stop here and do not apply 3D reconstruction algorithm. We proceed the obtained information in a different way. We warp image $g$ according to the established correspondence. The result denoted by $g'$, ideally, should be equal to image $f$. In reality, $g'$ and $f$ are not the same due to missing areas due to occlusions, errors in registration of pixels, change of illumination, small errors in interpolation, and *properties of the material*. So then we compare $g'$ and $f$, and measure its difference as a matchability map which is defined in terms of scalar product of image gradients,

$$M(f,g) = \frac{(\nabla f) \cdot (\nabla g')}{\|\nabla f\| \|\nabla g'\|}. \tag{1}$$

We found that the matchability map can be used for fabric recognition. More precisely, we found that fluffy surfaces are less matchable then smooth surfaces.

### B. Illustration of the method

An input of the algorithm is a couple of images $f$ and $g$, taken from different points of view. In all our experiments, we took images from distance between 30 and 80 centimeters, and distance between the camera positions was from 10 to 20 centimeters. These dimensions were chosen to be similar to



Fig. 3. Given two images $f$ and $g$. The images depict three materials: white and black pieces of fabric and linoleum flooring.
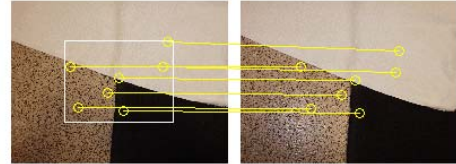


Fig. 4. Finding dense matching between images. Only a few matches are depicted here.

those which are typically used by a human, and convenient for robots. In our research, we took images of fabric and firm materials such as wood, linoleum, stone flooring, and paper. We explain and illustrate the algorithm with an example.

Obtaining the images is the first stage, so then we have two images in Fig. 3 as inputs for the algorithm.

In the second stage, we find correspondence between pixels in both images. In the first image, we delineate a rectangular area of interest, shown in Fig. 4, and we aim to obtain a matchability map in this area.

In the next stage, using the found correspondence, we warp the second image $g$ into a new image $g'$. Ideally, the two images, $f$ and $g'$ should coincide in the area of interest, as it is seen in Fig. 5. However, there are still differences between them and we assess the differences according to formula (1) and obtain the sought matchabilty map. The latter is present in the left image of Fig. 6.

The last stage is an interpretation of the matchabilty map. First of all, it is noisy and needs smoothing. After Gaussian smoothing we can clearly see that the map helps with recognition of fabric. Thresholding the smoothed map (with threshold=0.4) yields the result in the last of three images in Fig. 6. Here we can see that fabric was segmented out properly.

### C. Image registration method used in the presented experiments

To make the whole approach work, we need a well defined registration method, and we adopt an idea of using gradient and its orientation, see [1]. In [3], an "Orientation correlation, a fast, exhaustive, illumination invariant, statistically robust, translational image matching technique has been presented." A. Fitch [3] proved that matching orientations of gradient field leads to a robust method according strict definition of robustness. We chose this method for our purposes. However, taken as it is, this method is based on FFT and can only determine a pure translation. Thus we just take similarity

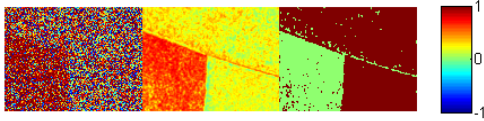Fig. 5. Warping. The first image is cropped $f$ and the second is a warped version of $g$.



Fig. 6. Matchabilty map, its smoothed version, and resulting segmentation of fabric

measure from [3], which is exactly the formula given in(1). To make it suitable for following curvature of the surface, coping with perspective transform, we use the similarity measure in conjunction with the registration method described by A. Averbuch and Y. Keller in [5].

Furthermore, we want to determine an affine transformation locally for each pixel $\vec{x} = (x_1, x_2)$, where each pixel belongs to a small patch $F_1$ in the image $f$. The affine transform $A$ is defined by

$$A(x_1, x_2) =$$

$$\left((1+\alpha_{11})x_1 + \alpha_{12}x_2 + \beta_1, \quad \alpha_{21}x_1 + (1+\alpha_{22})x_2 + \beta_2\right). \quad (2)$$

Now, we need to determine 6 parameters.

Let the patches $F_1$ and $F_2$ be taken from two images presenting the same part of the material, so then they are approximately equal. Now we want to refine the correspondence.

Gradient methodology [5] for the case of finding an affine transform $A^*$ is described as

$$A^* = \arg\min \sum_{\vec{x} \in F_1} (f(\vec{x}) - g(A(\vec{x}))^2 \quad (3)$$

which is then undergoes linearization procedure and iteratively solved.

Instead of intensity disparities [5] we use similarity measure (1), we modify (3) to the following,

$$A^* = \arg\max \sum_{\vec{x} \in F_1} \frac{(\nabla f(\vec{x})) \cdot (\nabla g'(A(\vec{x})))}{\|\nabla f(\vec{x})\| \|\nabla g'(A(\vec{x})\|}. \quad (4)$$

Then we apply bidirectional formulation of [5], and then we just simply follow the procedures in the cited paper to find $A^*$.
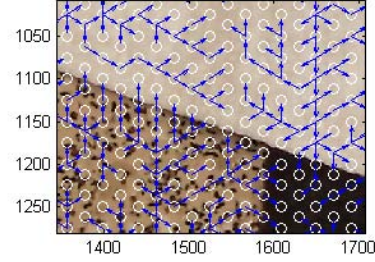


Fig. 7. Region growing, central part of the picture.

After the above explanation, we can describe the registration algorithm. The first step of the algorithm is to determine a couple of corresponding patches $F_1$ and $F_2$. It is done by the standard phase correlation method, which is applied to a chosen patch $F_1$. Then, according to the methodology explained above, an affine transform connecting patches $F_1$ and $F_2$ is obtained. At this point, we have one patch in the image $f$ with found affine correspondence. Then the concept of region growing is employed. All the patches have their centers in a triangle net, see Fig. 7.

Those patches can intersect, so we count them by their centers. Some patches already have found their correspondence, and they comprise a region to grow by accepting adjacent patches. Each time, a question arises, namely, which patch is the next to accept. For an answer, first we find the best patch among those located in the border of the region. The word "best" means best matching its counterpart in the second image. Thus the region grows by taking a new patch adjacent to the best patch. Relations between the best patch and its adjacent one are depicted in Fig. 7 by arrows. The new patch already has an approximate estimation of its affine transform due to proximity of the "best" patch, so there is no need to use the phase correlation method any longer. This procedure continues until the region of interest is fully registered.

The parameters we used are listed as follows,

- Image size 2000×3000; images are converted to gray level images;
- Images are Gaussian filtered with $\sigma = 2.2$ for better estimating for gradients;
- Patch size is 75×75;
- Nearest distance between centers of patches is 25.

## IV. EXPERIMENTAL VALIDATION

We performed a few experiments similar to the one described in Section III-B with different fluffy fabric and different smooth textured backgrounds. Each time we got similar result as in Fig. 6. In this section we present experiments with a special meaning.

### A. A stair case: two stair steps

Suppose a robot climbs on a staircase (Fig. 8) and watches two stair steps, Fig. 9. Please notice that the images have slightly different color, and it is normal because the images have to taken from different positions; however, this does
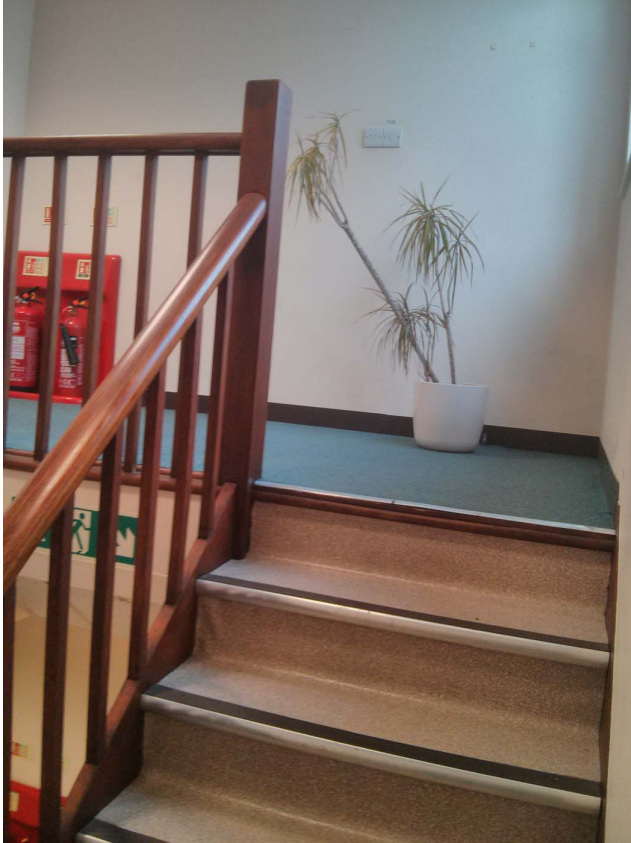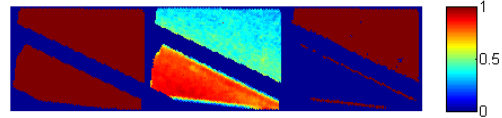
Fig. 10. **Two stair steps.** From left to right images depict: two regions of interest, matchability map, segmenation of fabric based on the map.



Fig. 11. Scene of two real fabric pieces are shown in the lower two quadrants. Their flat photographs are placed in the upper two quadrants.

piece of fabric is real or it is just a high quality picture of the given fabric; and nowadays warping complicated textures on flooring and other surfaces is common. Is there a way to discriminate them? Here we show that it is possible with the matchability map. This, taken alone, already justifies the offered approach.

The scene in Fig. 11 shows two pieces of fabric in the lower half of the image and a *photograph of the same* pair of the fabric pieces in its upper part.

The fabric pieces are colored, but in the photograph colors are absent due to an available printer. However, color does not participate in the task, because in the next step we transform images to greyscale digital images.

Then the task appears with the data in the first two greyscale images of Fig. 12 and it is to recognize the true fabric pieces. We applied the proposed approach and get segmentation result in the third image in Fig. fig:fabricPaper2. Therefore, the task is successfully solved.

## V. Discussion and future work

Positive features of the method are:

- There is no need in geometrical particularities, namely the positions of the camera and distance to the surface are not used;
- No previous data is required, no learning is involved.
- The material can move between photoshots, so then there is no need in a rigid scene.

Due to the construction of the method, some limitations are natural:

- The involved images should be of high resolution, and this affects time of computation;
- All the photo-pictured materials should have a texture otherwise one cannot match them;

Generally, the method as it, is effective only if "fluffiness" is prominent, otherwise, as it seems, much better resolution would be required. This is why it the presented method should be used in conjunction with other methods to enhance them.



Fig. 8. Staircase



Fig. 9. **Two stair steps.** The upper step is covered by a carpet, and the lower one is made of linoleum.

not influence the result due to the chosen algorithm. In this situation robot can be interested in haptic properties of the steps, since material of flooring may affect movement of the robot. Assuming that the robot is intelligent and it can recognize that there are two different materials and it can specify two regions of its interest. In Fig. 10 the two regions are shown in the left image.

Then, the robot can follow the proposed algorithm and compute matchability map $M(f, g)$, which is depicted in the middle Fig. 10. This map clearly says that the materials are different. Further thresholding gives the right image in Fig. 10, making a tentative suggestion that the upper step is covered by fabric.

### B. Fabric vs. its photograph

This is a more sophisticated example. In the Introduction, we noticed that existing methods cannot determine whether a

Fig. 12. The task is to recognize true fabric. The first two images comprise and input of the algorithm. In the third image a solution is given.

Let us consider potential developments. In the presented experiments, we place fabric and the background in the same image. It might be necessary for a human to make a comparison, while it is not true for a robot, since a robot can store pictures precisely and make comparison between new pictures with stored ones. This can be explored in future. Moreover, from a pair of images, which we already analysed in this paper, a robot can reconstruct a position of the surface and the position of the camera. Then the robot will be able to compare images stored in its memory with new acquired images taking into account the true angles of view; and the robot will be able to gradually build its own database accompanied with its haptic experience. Thus, the method grants a "sense of material" to a robot.

Here we would like to underline that the purpose of this conference paper is only to announce and describe the idea and to show that it works.

The following is not of the paper and is aimed for further work:

- to develop theoretical background;
- to conduct many experiments to behaviour of different materials;
- to extend the topic beyond smooth/fluffy materials;
- to try different matching algorithm.

## CONCLUSION

We presented a new idea for surface characterization based on the discovery that images of a patch of a flat surface are more correlated to each other than images of a patch, which belongs to a fluffy surface. This was explained and demonstrated in experiments. A matchability map was introduced to present the quantitative information of such correlation. We also presented an algorithm to calculate the matchability map and corresponding procedures to interpret it.

In a special experiment, we distinguished real fabric from its photograph by the proposed method. This is an important task because nowadays *any* textures are artificially mapped on surfaces, especially on plastic and paper surfaces, and to our knowledge existing methods in principle cannot distinguish fabric from its photograph.

We discussed limitations of the method and future work in this direction and concluded that the method can add reliability to other methods of segmentation, fabric recognition, material classification, estimating haptic properties of objects by robots.

## REFERENCES

[1] Thomas Brox and Jitendra Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(3):500–513, 2011.
[2] Oana G. Cula and Kristin J. Dana. 3d texture recognition using bidirectional feature histograms. *Int. J. Comput. Vision*, 59(1):33–60, 2004.
[3] A. J. Fitch, A. Kadyrov, W. J. Christmas, and J. Kittler. Orientation correlation. In *British Machine Vision Conference*, pages 133–142, 2002.
[4] B. Julesz. Visual pattern discrimination. *IRE Trans. Information Theory*, 8:84–92, 1962.
[5] Y. Keller and A. Averbuch. Fast motion estimation using bidirectional gradient methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(8):1042–1054, 2004.
[6] L.Liu and P.W.Fieguth. Texture classification from random features. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(3):574–586, 2012.
[7] T. Maenpaa M. Pietikainen, T. Nurmela and M. Turtinen. View-based recognition of 3d-textured surfaces. In *Proceedings 12th International Conference on Image Analysis and Processing. 17-19 Sept*, pages 530–535, 2003.
[8] Dimitra Triantafyllou and Nikos A. Aspragathos. A vision system for the unfolding of highly non-rigid objects on a table by one manipulator. *Intelligent Robotics and Applications, Lecture Notes in Computer Science.*, 7101:509–519, 2011.
[9] Manik Varma and Andrew Zisserman. A statistical approach to material classification using image patch exemplars. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(11):2032–2047, 2009.