

THE KTH-TIPS2 database

P. Mallikarjuna, Alireza Tavakoli Targhi, Mario Fritz,
Eric Hayman, Barbara Caputo and Jan-Olof Eklundh

Computational Vision and Active Perception Laboratory (CVAP)
School of Computer Science and Communication
KTH, SE-100 44 Stockholm, Sweden
{att,hayman, caputo, joe}@nada.kth.se

June 9, 2006

1 Background

This document provides a brief Users' Guide to the KTH-TIPS2 image database (KTH is the abbreviation of our university, and TIPS stands for Textures under varying Illumination, Pose and Scale). The KTH-TIPS2 provides a considerable extension to our previous database of images of materials — KTH-TIPS.

The guide describes which materials are contained in the database (Section 2), how images were acquired (Section 3) and subsequently cropped to remove the background (Section 4), and we also discuss some non-ideal artifacts, like poor focus, in some pictures (Section 5).

The objective with this database was to provide a more satisfactory means of evaluating algorithms for classifying materials. As we argued in [4, 1], a very relevant task is to recognise *categories* of materials such as “wood” or “wool” as opposed to one particular physical sample. The KTH-TIPS2 contains four physical samples of 11 different materials.

In addition, it is frequently necessary to perform recognition in unstructured environments. Thus the database provides images with variations in *scale* as well as variations in *pose* and *illumination*, following on from the philosophy of the KTH-TIPS, and in part the CURET image database [2].

The 11 materials in the KTH-TIPS2 database are all present also in the CURET database [2], which opens the possibility of conducting experiments on a combination of the two databases.

The cropped database is freely available on the internet [5]. Those interested in the full-size images should contact Eric Hayman (hayman@nada.kth.se).

The database was first presented and used in [1].

2 Imaged materials

The KTH-TIPS2 database contains images of 11 materials (Table 1 and Figure 1), each of which are also present in the CURET database [2], and six of which were also included in the first KTH-TIPS database [3]. Each of the samples is planar.

Material	Corresponding CURET sample number	Present in KTH-TIPS
Crumpled aluminium foil	15	×
Cork	16	
Wool	22	
Lettuce leaf	23	
Corduroy	42	×
Linen	44	×
Cotton	46	×
Brown bread	48	×
White bread	52	
Wood	54	
Cracker	59 and 60	×

Table 1: The materials present in the KTH-TIPS2 database.

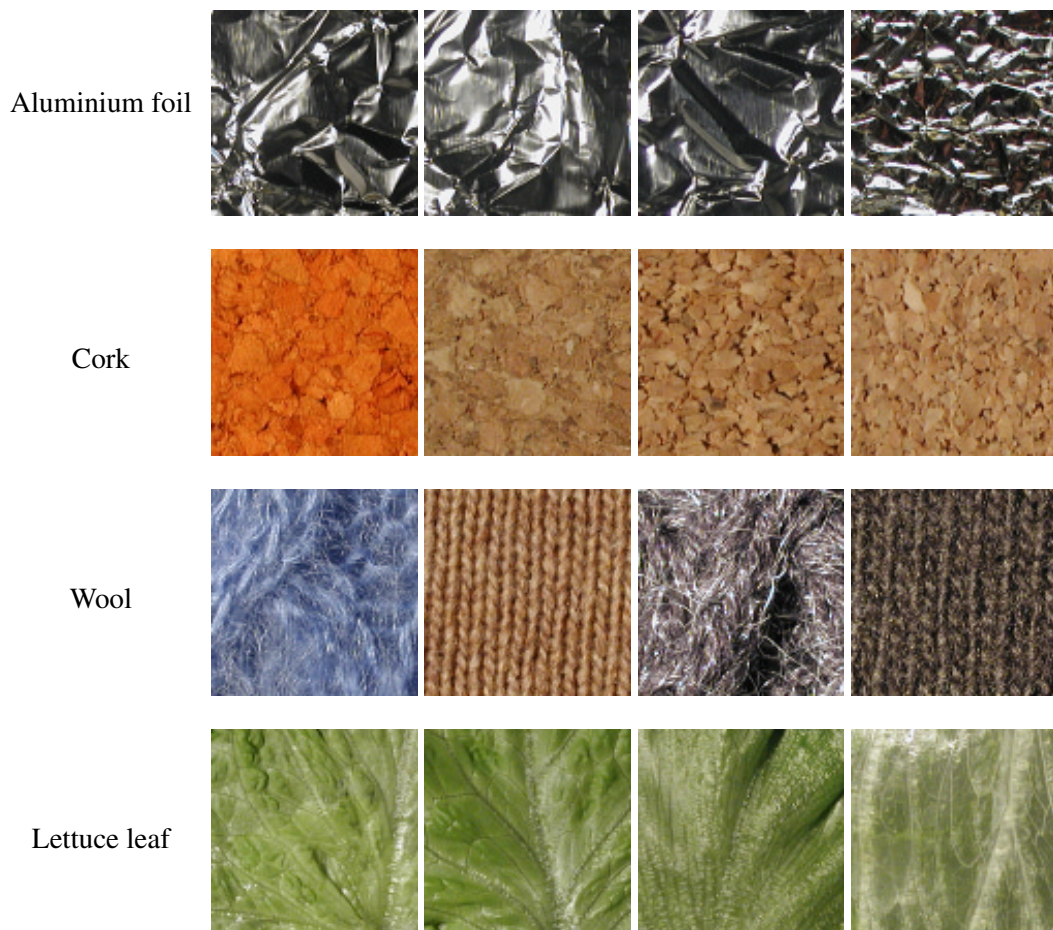


Figure 1: The variations within each category of the new TIPS2 database. Each row shows one example image from each of four samples of a category. In addition, each sample was imaged under varying pose, illumination and scale conditions. Continued on the next page.

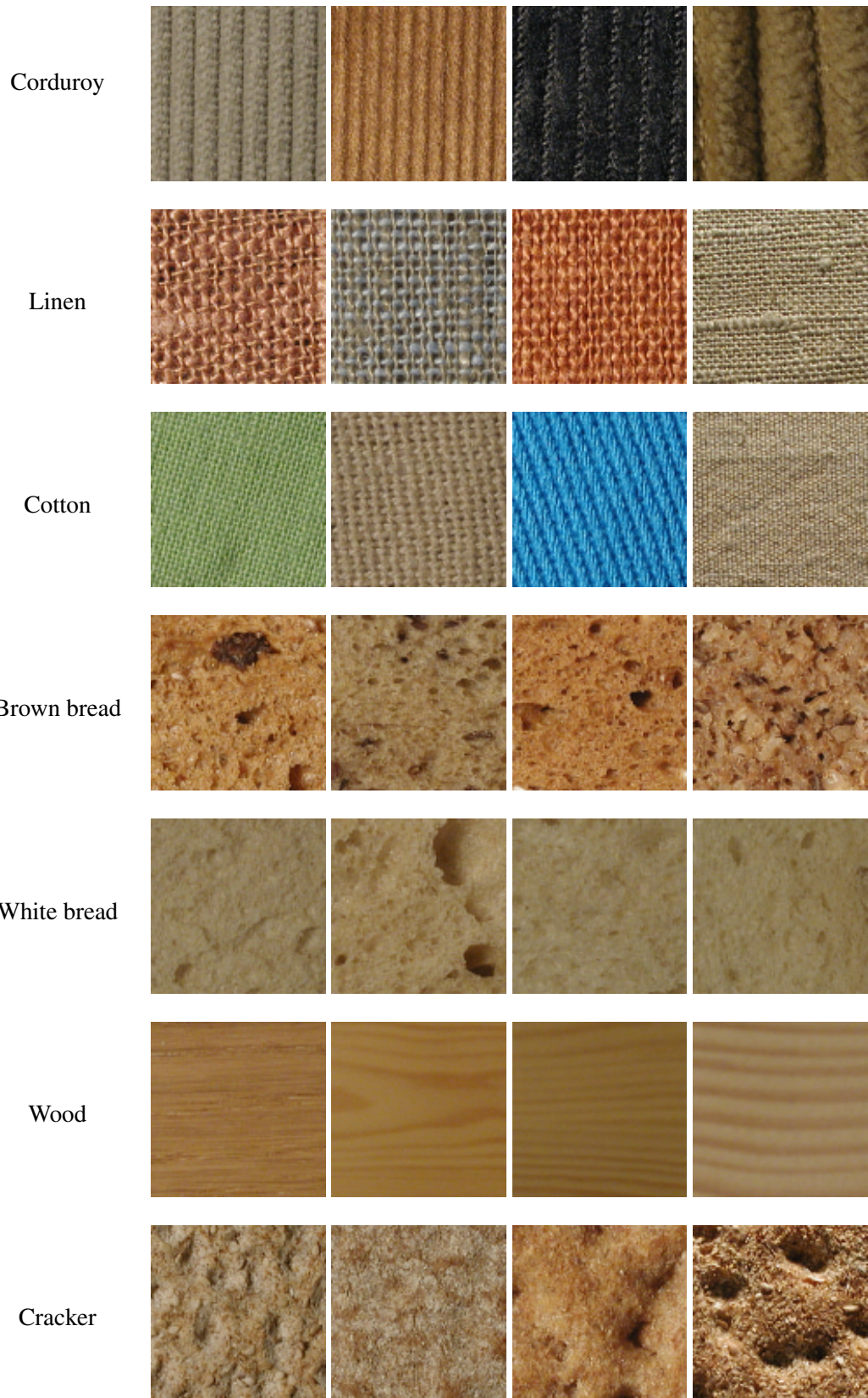


Figure 1 continued from the previous page.

The variation in appearance between the samples in each category is larger for some categories than others. Cork, for instance, contains relatively little intra-class variation, while cracker and wool exhibit significant variation. Note also that the appearance of wool depends not only on the material, but also on how it has been treated, in this case how the thread was spun and subsequently knitted. Moreover, it is not always obvious how the samples should be split into categories. For instance, brown bread and white bread are subclasses of “bread”, and it might also make sense to group linen and cotton together in a “woven fabric” class. As such we believe that this database provides a good platform for future studies of unsupervised or supervised grouping of classes into higher-level categories, whether visual or semantic, in a hierarchical structure.

3 Image acquisition

The acquisition of KTH-TIPS2 images largely followed the procedure used for KTH-TIPS as described in [3], though with some differences with regard to *scale* and *illumination*.

The images were taken with an Olympus C-3030ZOOM digital camera at a resolution of 1280×960 pixels. Many of the full-size images contain not only the sample, but also some background.

Like KTH-TIPS, KTH-TIPS2 contains images at 9 scales equally spaced logarithmically over two octaves. However, in KTH-TIPS2 the scale closest to the camera corresponds to Scale #2 of KTH-TIPS. This is due to problems with focus in Scale #1 of KTH-TIPS. The scales used are described in full in Table 4, and full-resolution images from one material (Cracker B) are shown in Figure 2. To maintain compatibility with KTH-TIPS, these scales are labelled with numbers ranging from 2 to 10 rather than 1 to 9. In other words, scale 4 in KTH-TIPS and scale 4 in KTH-TIPS2 represents the same camera-object distance.

KTH-TIPS2 contains images at the same 3 poses as KTH-TIPS (frontal, rotated 22.5° left and 22.5° right), but 4 rather than 3 illumination conditions. The 3 illuminations from KTH-TIPS are used (frontal, 45° from the top and 45° from the side, all taken with a desk-lamp with a Tungsten light bulb), and for the fourth illumination condition we switched on the fluorescent lights in the laboratory.

At each scale 12 images were taken in a combination of three poses (frontal, rotated 22.5° left and rotated 22.5° right) and four illumination conditions (from the front, from the side at roughly 45° and from the top at roughly 45° , and using ambient lighting). The adopted labelling scheme is shown in Table 5 and sample images in Figure 3. Note that images 1–9 follow the naming convention of KTH-TIPS, whereas images 10–12 are the images with the new, ambient, lighting condition.

This gives a total of $12 \times 9 = 108$ images per sample. **However:** For “Aluminium foil”, “Linen”, “Cotton” and “Cracker”, i.e. four of the six classes already in KTH-TIPS (Table 1), we first only acquired images from 3 new samples, and inserted the original KTH-TIPS images into the KTH-TIPS2 database as the fourth sample *with the caveat that those four samples were not captured with ambient lighting, and only 8 scales (2–9) were available.* This implies that in our first research using KTH-TIPS2 [1], for four physical samples out of 44, only $9 \times 8 = 72$ images were used. We have since taken further images, implying that we do now have 108 images for each sample, but we have not yet rerun our experiments from [1] using also these images.

We therefore define two versions of KTH-TIPS2:

- **KTH-TIPS2-a:** The version used in [1] with only 72 images used for 4 out of 44 samples.
- **KTH-TIPS2-b:** The version containing 108 images for all 44 samples.

Both versions are available for download from [5].

Scale number	Relative scale	Distance to camera (cm)
2	$2^{-1.00} = 0.500$	16.65
3	$2^{-0.75} = 0.595$	19.80
4	$2^{-0.50} = 0.707$	23.55
5	$2^{-0.25} = 0.841$	28.00
6	$2^{0.00} = 1.000$	33.30
7	$2^{+0.25} = 1.189$	39.60
8	$2^{+0.50} = 1.414$	47.09
9	$2^{+0.75} = 1.682$	56.00
10	$2^{+1.00} = 2.000$	64.41

Table 4: The scales present in the KTH-TIPS2 database.

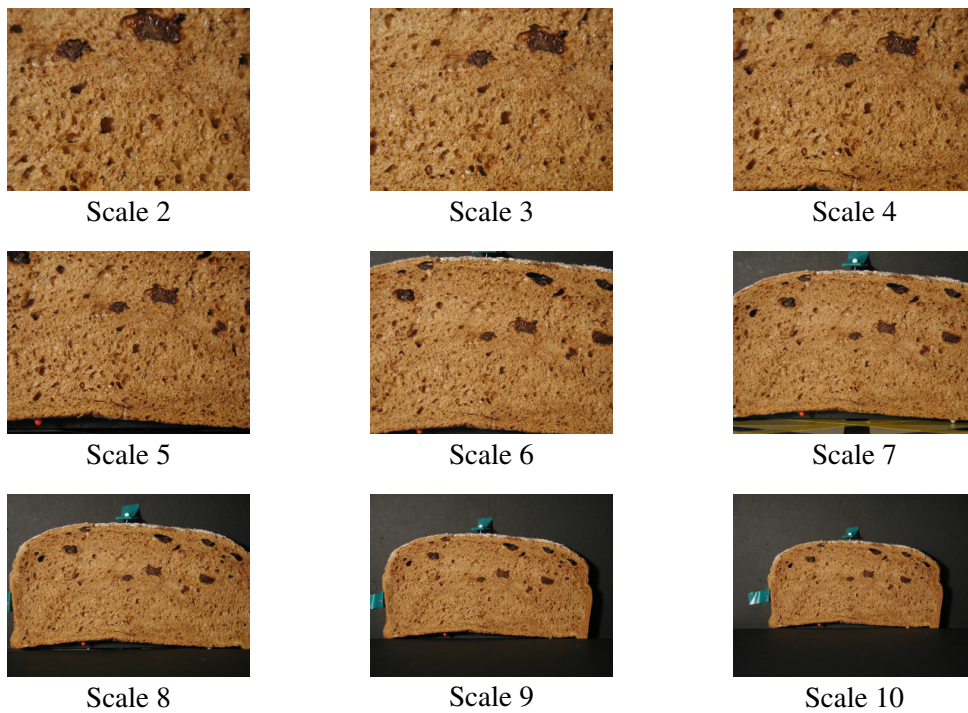


Figure 2: Full-size images depicting the variation of scale present in the KTH-TIPS2 database.

Image number	Object pose			Illumination direction			
	Frontal	22.5 ° right	22.5 ° left	Frontal	≈ 45 ° from top	≈ 45 ° from side	Ambient
1	x			x			
2	x				x		
3	x					x	
4		x		x			
5		x			x		
6		x				x	
7			x	x			
8			x		x		
9			x			x	
10	x						x
11		x					x
12			x				x

Table 5: The labelling of images within each scale in the KTH-TIPS2 database.

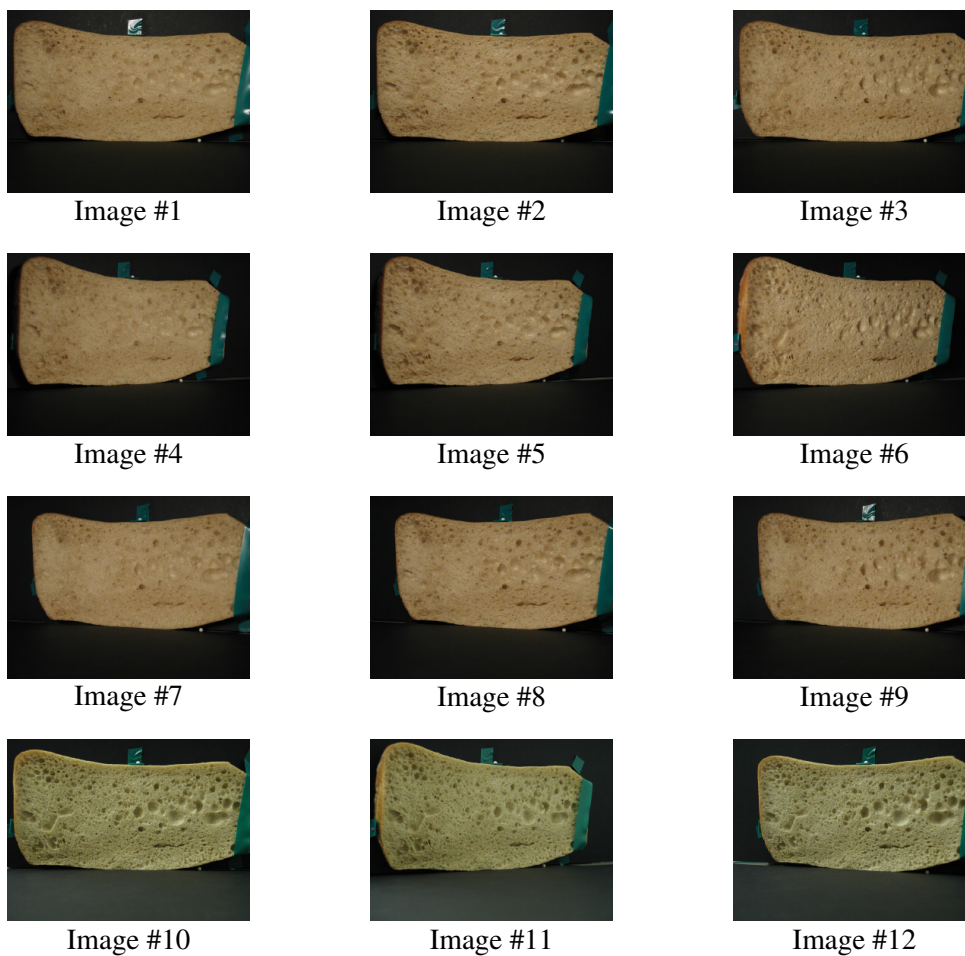


Figure 3: The variation of pose and illumination present in the KTH-TIPS2 database. In each row the pose is constant, whereas in each column the illumination is the same (frontal, side, or top illumination).

4 Image cropping

To remove the background, and to be consistent with the experiments in [6, 7], we manually cropped images to 200×200 pixels when possible. However, for some samples this was not possible at large camera-target distances since the subject did not fill a sufficiently large part of the image. This issue was also discussed in the documentation for KTH-TIPS [3], and a procedure for cropping the images was described there. We used the same procedure for KTH-TIPS2. In fact, this problem is greater with KTH-TIPS2 because we replaced Scale #1 with Scale #10 due to problems focusing. We note that this cropping was based on a notion of equivalent size in terms of the number of pixels available to our default subsequent image processing, and is not entirely fair when comparing different algorithms.

Table 6 and 7 list where these cropping strategies were necessary for KTH-TIPS2-a and KTH-TIPS2-b respectively. With Brown bread the texture round the edges of the slice is somewhat different (denser) to that in the middle, so these edges were also removed.

5 Some poor quality images in the database

As in KTH-TIPS, we again encountered some problems when acquiring images for KTH-TIPS2, implying that not all images were up to satisfactory standards. We have already discussed the issue of *cropping* and that our first version of the database, KTH-TIPS2-a, only contained 72 rather than 108 images for 4 out of 44 samples.

We also mentioned the problems associated with focusing at close distances. Despite removing Scale 1, we still have some images out of focus when the subject is close to the camera. Moreover, some images of fine-structured materials appear out of focus at larger distances since the images appear too homogeneous. Furthermore, in an attempt to get patches which were in focus, non-central patches were often selected, implying that the estimate of camera-target distance is inaccurate. Table 8 summarises the images with poor focus.

Further issues are

- perspective effects at some of the closer distances (this is fairly minor);
- minor creases in cloth (linen and cotton) induce spurious edges. On the other hand, real-world samples of cotton and linen can also have creases;

References

- [1] B. Caputo, E. Hayman, and P. Mallikarjuna. Class-specific material categorisation. In *Proc. Int. Conf. on Computer Vision*, pp II: 1597–1604, 2005.
- [2] K.J. Dana, B. van Ginneken, S.K. Nayar, and J.J. Koenderink. Reflectance and texture of real-world surfaces. *TG*, 18(1):1–34, January 1999.
- [3] M. Fritz, E. Hayman, B. Caputo, and J.-O. Eklundh. The KTH-TIPS database. Available at www.nada.kth.se/cvap/databases/kth-tips.
- [4] E. Hayman, B. Caputo, M. Fritz, and J.-O. Eklundh. On the significance of real-world conditions for material classification. In *Proc. 8th Eur. Conf. on Computer Vision, Prague*, pp IV:253–266, 2004.
- [5] P. Mallikarjuna, M. Fritz, A. Tavakoli Targhi, E. Hayman, B. Caputo, and J.-O. Eklundh. The KTH-TIPS and KTH-TIPS2 databases. Available at www.nada.kth.se/cvap/databases/kth-tips.
- [6] M. Varma and A. Zisserman. Classifying images of materials: Achieving viewpoint and illumination independence. In *Proc. 7th European Conf. on Computer Vision, Copenhagen*, pp III: 255 ff., 2002.
- [7] M. Varma and A. Zisserman. Texture classification: are filter banks necessary? In *Proc. Computer Vision and Pattern Recognition*, pp II: 691–698, 2003.

Material	Sample	Scale	Images	Cropping strategy
Cork	b	10	9	Largest possible
Lettuce leaf	a	7	1	Equivalent size
		9	12	Equivalent size
		10	3,4,5	Largest possible
		10	6,7,8,9	Equivalent size
Corduroy	b	7	6	Equivalent size
		8	9	Equivalent size
		9	4,5,6,7,9,10,11	Equivalent size
		10	2,3,4,5,8,9	Equivalent size
		10	10,11,12	Largest possible
	c	2	5	Equivalent size
		3	6	Equivalent size
		8	1,2	Equivalent size
		9	11	Equivalent size
Cotton	c	7	11	Equivalent size
		9	3	Equivalent size
		10	8,9	Equivalent size
Linen	b	10	12	Equivalent size
		4	4	Equivalent size
	c	7	12	Equivalent size
		8	2	Equivalent size
		10	4	Equivalent size
Brown bread	b	10	9,10	Equivalent size
		10	12	Largest possible
	c	10	1,6,11,12	Equivalent size
d	10	11,12	Equivalent size	
White bread	c	10	11	Equivalent size
		9	6,7,8	Equivalent size
	d	10	1-8,10,11,12	Equivalent size
		10	9	Largest possible
Wood	b	6	10	Equivalent size
		5	9	Equivalent size
	d	8	9,12	Equivalent size
Cracker	a	5	4,5,6,7,12	Equivalent size
		6	1-6, 8-12	Equivalent size
		7	1-8,10,11	Equivalent size
		7	9,12	Largest possible
		8,9,10	All	Largest possible
	b	6	5,9,12	Equivalent size
		7	1-3, 6-12	Equivalent size
		7	4,5	Largest possible
		8,9,10	All	Largest possible
	d	7	All	Equivalent size
		8	1,2,3	Equivalent size
		8	4-9	Largest possible
		9	All	Largest possible

Table 6: Images in KTH-TIPS2-a where it was not possible to extract 200×200 pixels foreground patches.

Material	Sample	Scale	Images	Cropping strategy
Cork	b	10	9	Largest possible
Lettuce leaf	a	7	1	Equivalent size
		9	12	Equivalent size
		10	3,4,5	Largest possible
		10	6,7,8,9	Equivalent size
Corduroy	b	7	6	Equivalent size
		8	9	Equivalent size
		9	4,5,6,7,9,10,11	Equivalent size
		10	2,3,4,5,8,9	Equivalent size
		10	10,11,12	Largest possible
	c	2	5	Equivalent size
		3	6	Equivalent size
		8	1,2	Equivalent size
		9	11	Equivalent size
Cotton	c	7	11	Equivalent size
		9	3	Equivalent size
		10	8,9	Equivalent size
Linen	b	10	12	Equivalent size
		4	4	Equivalent size
	c	7	12	Equivalent size
		8	2	Equivalent size
		10	4	Equivalent size
Brown bread	b	10	9,10	Equivalent size
		10	12	Largest possible
	c	10	1,6,11,12	Equivalent size
d	10	11,12	Equivalent size	
White bread	c	10	11	Equivalent size
		9	6,7,8	Equivalent size
	d	10	1–8,10,11,12	Equivalent size
		10	9	Largest possible
Wood	b	6	10	Equivalent size
		5	9	Equivalent size
	d	8	9,12	Equivalent size
Cracker	a	5	4,5,6,7,12	Equivalent size
		6	1–6, 8–12	Equivalent size
		7	1–8,10,11	Equivalent size
		7	9,12	Largest possible
		8,9,10	All	Largest possible
	b	6	5,9,12	Equivalent size
		7	1–3, 6–12	Equivalent size
		7	4,5	Largest possible
		8,9,10	All	Largest possible
	d	7	1–9,11,12	Equivalent size
		8	1,2,3,10,11,12	Equivalent size
		8	4–9	Largest possible
		9,10	All	Largest possible

Table 7: Images in KTH-TIPS2-b where it was not possible to extract 200×200 pixels foreground patches. Note that the only difference relative to KTH-TIPS2-a (Table 6) is in Cracker, sample d.

Material	Sample	Scale	Images
Cork	a	2	1-8
	c	4	11
		10	5
	d	2	7,8
10		1,4,5,6,11	
Wool	b	2	7,8,9
	c	5	6,10
	d	2	10
		9	11
10	10,11,12		
Lettuce leaf	a	2	7
	d	2	1-6,10,11
Corduroy	b	5	11
		6	2
	c	8	1-6,10,11
		9	4,5,6
10	6		
Cotton	a	2	2,7,10
		8	7,8,9,11,12
		9	4-9,11,12
	b	2	12
	d	8	8
9		1,4,7,8,9	
Brown bread	b	2	1,2,3
White bread	d	2	7,8,9,12
Wood	d	3	11
		5	4-9,12
Cracker	b	2	1

Table 8: Images poorly focused in the KTH-TIPS2-a and KTH-TIPS2-b databases. None of the images added to KTH-TIPS2-a to form KTH-TIPS2-b suffered from blur.