

Journal of Electronic Imaging

JElectronicImaging.org

Material recognition by feature classification using time-of-flight camera

Fabio Martino
Cosimo Patruno
Nicola Mosca
Ettore Stella



Fabio Martino, Cosimo Patruno, Nicola Mosca, Ettore Stella, "Material recognition by feature classification using time-of-flight camera," *J. Electron. Imaging* **25**(6), 061412 (2016), doi: 10.1117/1.JEI.25.6.061412.

Material recognition by feature classification using time-of-flight camera

Fabio Martino,* Cosimo Patruno, Nicola Mosca, and Ettore Stella

National Research Council of Italy, Institute of Intelligent Systems for Automation (CNR-ISSIA), via Amendola 122/DO, Bari 70126, Italy

Abstract. We propose a method for solving one of the significant open issues in computer vision: material recognition. A time-of-flight range camera has been employed to analyze the characteristics of different materials. Starting from the information returned by the depth sensor, different features of interest have been extracted using transforms such as Fourier, discrete cosine, Hilbert, chirp-z, and Karhunen–Loève. Such features have been used to build a training and a validation set useful to feed a classifier (J48) able to accomplish the material recognition step. The effectiveness of the proposed methodology has been experimentally tested. Good predictive accuracies of materials have been obtained. Moreover, experiments have shown that the combination of multiple transforms increases the robustness and reliability of the computed features, although the shutter value can heavily affect the prediction rates. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JEI.25.6.061412]

Keywords: material recognition; material classification; time-of-flight sensor; predictive machine-learning model.

Paper 16203SS received Mar. 15, 2016; accepted for publication Aug. 2, 2016; published online Aug. 23, 2016.

1 Introduction

The material of an object can be considered significant data for understanding one or more scenes. We usually interact with a wide variety of materials, and we continually assess their properties such as weight, size, and texture. Knowledge of such properties could be useful for a robot manipulator that has to handle an extensive assortment of objects. As an example, the physical property can be advantageous to discern breakable objects from the robust ones. In this way, a manipulator gripper force can be tuned to avoid item damage.

The material information can also be useful in other applicative contexts such as robot localization and environmental mapping where three-dimensional (3-D) data are employed. In this regard, a better 3-D point cloud registration could be achieved by exploiting knowledge of the material type. As an example, a complex environment made of glass and highly reflective surfaces could be more challenging for registration algorithms. Therefore, a preliminary analysis aimed to identify the material type would be helpful for discarding some 3-D points from the registration method. In this way, only the stable points referring to nonchallenging materials are considered in the computation, implying an enhancement of registration accuracy.

Nevertheless, material recognition is currently a difficult challenge. Over the years, many works have been proposed to achieve material classification.^{1–6} Most of the common approaches are based on color analysis or textural appearance. In this regard, a method based on 3-D textons¹ was introduced to recognize surfaces on the basis of their textural appearance. A vocabulary of tiny surface patches together with their local and photometric properties was built to characterize the local irradiant distribution.

Other textural representations² based on fast Markovian statistics were proposed for recognizing natural materials. The proposed features are fast to compute and robust to illumination direction as well as invariant to brightness changes. A good predictive accuracy was achieved from the analysis of several natural materials acquired under varying view-points, illumination colors, and directions.

A rich set of local features³ exploiting the Kernel descriptor framework combined with large-margin nearest neighbor learning was empirically studied to accomplish the material recognition of real-world objects as well.

A method that exploits several features covering various aspects of material appearance was also proposed for material classification.⁴ The support vector machine (SVM) framework was employed for obtaining a recognition rate of 53.1%, much better than the predictive rate obtained by using the Bayesian inference framework.⁷

High predictive accuracies were also achieved by using the algorithm proposed in Ref. 5. Specifically, the material appearance is modeled as the joint probability distribution of responses extracted from filter bank and color values (in the hue-saturation-value space). SVM was then employed as a classifier. By considering image patches of resolution 30×30 pixels, a very high accuracy was found.

A framework called reflectance hashing⁶ was introduced to model the reflectance disk of a material surface acquired from a unique optical camera measuring technique. The high-dimensional reflectance is encoded with a compact binary code that efficiently reveals the material class.

Liu et al.⁷ and Sharan et al.⁸ computed different low- and middle-level features to assess the appearance of materials. Then an augmented latent Dirichlet allocation (aLDA) method based on a Bayesian framework was applied to combine such features.

Large-scale datasets combined with deep learning were also proposed for scene classification and object recognition.^{9,10}

*Address all correspondence to: Fabio Martino, E-mail: martino@ba.issia.cnr.it

In detail, a convolutional neural network (CNN) was employed to classify the materials. A mean recognition rate of about 80% was obtained.

An alternative statistical approach was presented in Ref. 11, where the joint distribution of intensity values of single images was employed together with filter banks providing state-of-the-art classification rates.

Most of the techniques presented in the literature for material recognition exploit passive 2-D cameras. However, the reflectance properties of material, the kind of surface (smooth or rough), the illumination, and the view angle conditions could compromise material identification. These cameras always need illumination in dark environments, and lighting variations could significantly complicate a real-scene analysis.

Nevertheless, the recent design of 3-D range sensors has gained significant interest for a wide number of applications.^{12,13} In fact, most issues related to 2-D cameras in material recognition tasks can be overcome by using time-of-flight (ToF) cameras. As an example, it is not necessary to use an external light source because such acquisition systems are able to sense the neighboring environment by employing infrared (IR) light. Hence, useful information can be collected even when the objects to be examined are poorly lit. Moreover, the 3-D data returned by ToF sensors provide significant information about the geometry and the shape of objects located in a scene. Therefore, problems tied to both view angle conditions and roughness of surfaces are considerably reduced when a 3-D sensor is used instead of 2-D ones.

Although ToF cameras can be employed only in indoor environments where the sunlight cannot interfere, the other benefits gained by employing these sensors enable better investigation of the material properties to accomplish material classification. This paper will deal with methodology for material recognition by exploiting a ToF range camera.

Similar works that exploit 3-D sensors and noncontact active techniques¹⁴⁻¹⁶ were presented to evaluate the object material. Specifically, the geometric properties of a material were investigated through the analysis of the reflected pattern of IR light. Microstructural details of materials and other associated information, i.e., shape and color, were computed by utilizing a ToF camera. The patterns related to materials were then classified by a random forest (RF) classifier.

This paper presents an alternative technique for achieving material recognition exploiting the data given from a ToF camera. The basic idea is to analyze whether a correlation can be established between the type of material of an object and the alterations affecting measurements taken with a ToF sensor. For every tested material, a patch from the 3-D point cloud dataset is extracted. At this stage, features based on different domains of transform (e.g., discrete cosine transform, Fourier transform, Hilbert transform, and so on) are computed to characterize the material.

Several working conditions have been taken into account, such as the pose of the material with respect to the depth sensor and the shutter value of the photoreceivers. A decision tree (J48) has then been employed to classify the materials.

This paper is organized as follows: some important aspects regarding ToF sensors are discussed in Sec. 2, and the methodology employed for material recognition is

presented in Sec. 3. Experimental results and related discussion are reported in Sec. 4. Final conclusions and remarks are in Sec. 5.

2 ToF Range Camera: Depth Measurement Errors

As mentioned, the aim of this work is to identify the material category of an item (e.g., wood, metal, plastic, glass, fabric, and so on) by analyzing the information given from a range camera. This sensor exploits the well-known principle of ToF to profile the surrounding environment. Therefore, our main idea is to investigate the materials by analyzing the alterations affecting the measurements over time. In this regard, some physical material properties such as reflectance, scattering, and absorption might affect the IR light source of the ToF sensor that strikes the surface by involving fluctuations of returned information.

Since we take advantage of depth measurement alterations to accomplish material recognition, it is worth having a discussion about the unavoidable errors^{17,18} that might affect ToF sensors. In this regard, two main categories can be identified: systematic and nonsystematic depth errors. Typical systematic errors can be due to depth distortions (i.e., when an incorrect sinusoid is generated), lens distortions, integration time, operating temperature, overexposed reflected amplitudes, ambient light conditions, and so on. In contrast, the most common nonsystematic errors are due to multiple light reception, motion blurring, light scattering, signal-to-noise ratio (SNR) distortion, and so on.

Essentially, it is important to reduce the effect of these errors to ensure reliable material recognition. In fact, some precautions and compensation methods could be adopted to achieve our purpose.

Systematic depth measurement errors are mainly due to IR sinusoidal generators, which have limits in their modulation process. Such irregularities involve a phase perturbation caused by erroneous wrapping due to the presence of odd harmonics. Consequently, a change of depth value occurs, compromising the actual computation of distance. In the same way, other errors due to integration time and operating temperature can consistently affect the actual computation of the depth map as well. It is important to define a model of error to get more accurate and reliable depth measurements.¹⁹⁻²¹ In this regard, some details will be provided in Sec. 4, where the counter-measures adopted to limit these errors have been explained.

Another important aspect that has to be faced is the effect of lens distortion. Such effects are mainly due to the curvature of lenses mounted by ToF cameras. Therefore, precautionary steps have to be performed to decrease the distortions that affect the depth image. More details concerning the rectification step will be provided in Sec. 3.1.

The proper functioning of ToF cameras is strictly linked to ambient light conditions as well. In this regard, external waves having a comparable wavelength to the light source employed by the sensors for scanning the environment can compromise the reliability of measurements. Therefore, such sensors could be used only in indoor environments, as stated. Further details about the ambient light conditions of our tests will be reported in Sec. 4.

Other nonsystematic errors that might negatively affect measurements are mainly due to multiple paths of the light source, low SNR ratio, light scattering, and so on. Usually,

these errors can be managed by employing filtering methods or suitable models. Specifically, many works have been presented to investigate the effect of scattering of surfaces. Most of them have presented models based on the bidirectional reflectance distribution function.²² A more accurate model of light scattering was introduced in Ref. 23, where the bidirectional scattering-surface reflectance distribution function was described. Nevertheless, this model can be employed only to measure the scattering properties of translucent materials.

In this regard, some experiments will be conducted to examine in depth how scattering, reflectance, and absorption affect our methodology in material recognition. In other words, this paper will mainly focus on the description of our approach along with related results. Therefore, all physical aspects behind our idea will be addressed in future work.

3 System Overview

In this work, we have taken advantage of the reflectance and absorption of the material surfaces considering several working conditions. A ToF range camera has been employed to create the datasets for our experiments. Therefore, exploiting the 3-D information given by the sensor together with the related intensity values, a variety of materials has been investigated. Specifically, different features have been extracted and then evaluated by a decision tree to accomplish the material classification.

3.1 Mathematical Framework

Several mathematical methods have been compared to obtain a reliable classification of material type, such as:

- fast Fourier transform (FOURIER);
- discrete Hilbert transform (HILBERT);
- discrete cosine transform (DCT);
- Karhunen–Loève transform (KLT);
- chirp-z transform (CHIRP).

These transforms are commonly employed in signal processing tasks. Each of them has particular properties or characteristics that can aid the material analysis process. Since the material to be recognized is examined by analyzing several sequences of frames over time, it is important to use transforms that are able to extract significant features from signals.

In our tests, we have taken advantage of the fast Fourier transform (FFT), i.e., a faster alternative to discrete Fourier transform (DFT). In general, the FFT is a powerful tool for pattern recognition. It is commonly employed to extract invariant features²⁴ because of its important properties; for example, a shift in the time domain does not involve any change in the amplitude spectrum of the image. Good predictive accuracies in material recognition are expected since the frequency domain representation might provide more useful information than the time domain one. Moreover, a low processing time should be required for computing this transform.

The Hilbert transform extends a differentiable real signal into the Gauss plane. This transform adds information to the Fourier analysis because it introduces the conjugate harmonic of a given signal. Usually, it is used to handle

nonstationary processes or signals for which the Fourier spectral analysis is not often suited. In our algorithm, the discrete version is exploited.

The DCT is similar to the DFT since they both decompose a discrete-time signal into a sum of scaled and shifted basis functions. However, the DCT uses only cosine functions as its kernel. The DCT is widely used in signal processing and data compression applications because of its high-compression degree of spectrum. One of the most relevant properties^{25,26} is the noise high-frequency isolation in a small number of coefficients compared to other transforms such as the DFT. Average recognition rates are expected since the high-compression level of this transform might negatively affect the informative content of input signals.

The KLT is a representation of orthogonal functions. It has different expansion bases that depend on the stochastic process, and their coefficients are random variables. The kernels employed in this representation are defined by the covariance function of the process. Although the KLT has a high computational complexity, it is suitable to obtain the best bases for linear decorrelation of signals and energy compression. Hence, good predictive rates should be achieved by employing this mathematical tool.

Finally, the last employed transform is the chirp-z-transform (CZT), which can be considered a generalized case of DFT. In fact, the CZT samples the Z-plane along spiral arcs, which correspond to straight lines in the Laplace-plane. The kernel of this transform is a complex number.

All listed mathematical domains have been used and compared to extract features suitable for the material recognition method. Different materials having different physical characteristics have been examined. Particularly, the material target under investigation has been fastened on a panel and then placed in front of the acquisition system. Several working conditions have been taken into account. For instance, the shutter value (or exposure time), the position (or distance), and the angle (or heading) of panels with respect to the sensor have been varied.

It is necessary to emphasize that only a portion of the scene has been considered. Specifically, a region of interest (RoI) of the panel has been extracted, as reported in Fig. 1. The target has been centered in the middle of field-of-view (FoV) of the ToF sensor. In this way, the distortion effects due to the curvature of lenses have been consistently reduced. Nevertheless, a preliminary step aimed to rectify the depth images has been performed to compensate the distortion effects. In this regard, the camera calibration toolbox for MATLAB,²⁷ along with the well-known notation introduced

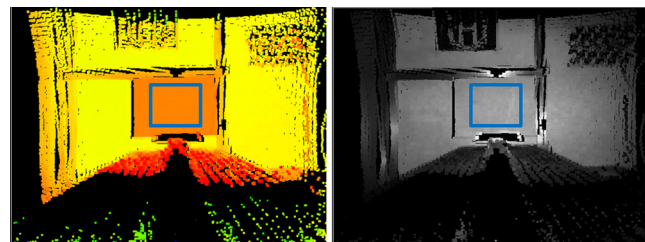


Fig. 1 Example of the extracted RoI of 20×25 pixels (see cyan rectangle) of a wooden panel during an experiment. The 3-D point clouds mapped according to both (a) distances and (b) IR reflectivity are shown.

by Heikkilä,²⁸ has been exploited to get a faithful reconstruction of scenes without distortions.

At this stage, a punctual analysis of the panel is performed. In other words, every pixel of the extracted RoI is evaluated over time. A sequence of n frames is acquired for every material. Thus, each pixel has n values of 3-D coordinates together with the related intensity levels

$$p(u, v) = \{s_1, s_2, \dots, s_k, \dots, s_n, i_1, i_2, \dots, i_k, \dots, i_n\}$$

$$k = 1, \dots, n, \quad (1)$$

$$s_k = d_k - \bar{d}, \quad (2)$$

$$d_k = \sqrt{X_k^2 + Y_k^2 + Z_k^2}. \quad (3)$$

Eq. (1) represents exactly how the data linked to a pixel of 2-D coordinates (u, v) is arranged. In this regard, the vector $p(u, v)$ contains both the distance variations s_k of the pixel and the associated intensity values i_k . Distance fluctuations are computed using Eq. (2). Specifically, the value of distance d_k is evaluated by means of Eq. (3), whereas \bar{d} represents the average value of all n distances related to the pixel under investigation. The terms X_k , Y_k , and Z_k represent the 3-D coordinates of the pixel (u, v) at the time k .

Subtracting the average distance from the set of measures fulfills a precautionary step aimed at reducing the possibility that material recognition might be achieved by the classifier by considering the evaluated distance instead of its fluctuations. Artificial biases introduced by the experiment itself in the material recognition process are therefore drastically decreased.

Once the input vector has been collected, the different transformation domains listed before have been computed. In this way, one or more features are associated to each pixel of the RoI.

3.2 Description of Decision Tree

As stated before, the aim of this paper is to provide a method able to discern the category of materials by analyzing the data returned from a ToF camera.

Material recognition is achieved by exploiting a decision tree as a model of classification. In brief, a decision tree is a predictive machine-learning model able to provide an output value by evaluating numerous attribute values of the available data. It can be considered a treelike graph that has nodes and branches. Specifically, the internal nodes denote the different attributes, whereas the branches between nodes specify the possible values these attributes can have. Finally,

the terminal nodes identify the final value or the classification output of the dependent variable.

Regarding this specific case, the attributes of the decision tree are the coefficients of the various computed features together with other parameters that characterize the performed experiment. Further information will be provided in Sec. 4.

In our tests, we employed the open source software Weka,²⁹ a collection of machine-learning algorithms, for fulfilling material recognition. Specifically, we chose the J48 classifier, a high-performance algorithm suitable for large datasets, which is a popular open-source implementation of the C4.5 decision tree (Ref. 30) classifier that is available in Weka. The employed classifier supplies an easy interpretation model and it is suitable for datasets, which are heavily affected by noise. Although other classifiers (e.g., RF) provide comparable predictive accuracies with respect to J48, they require higher computational processing. Moreover, preference was given to deterministic algorithms for repeatability reasons.

3.3 Data Arrangement

Section 3.1 reports the main domains of analysis for associating features to a single pixel under examination. However, material recognition is achieved by considering the evaluation of a set of pixels, i.e., those pixels belonging to the considered RoI.

As previously stated, a decision tree is designed to provide the category of material under examination. Therefore, the data of interest have to be properly arranged to be managed by the decision graph for material classification. In other words, starting from the computed features for a set of materials and considering other parameters tied to the typology of experiments, a suitable representation has been obtained and stored for processing using Weka. Specifically, the file has been organized according to the attribute-relation file format (ARFF).

The header of an ARFF file is reported in Table 1. The coefficients of the features are arranged along the columns, then other parameters such as position, angle, active brightness, and shutter are organized in the same way. The position and the angle values represent the pose of the panel to be examined with respect to the ToF sensor. Further details will be provided in the following section. The active brightness is a boolean parameter of the camera that enables improvement of the quality of acquisition. In fact, when this parameter is disabled, all measurements returned by the photoreceivers are considered good, even if some reflected light is captured by the photosensors. Therefore, it is preferable to enable this modality to obtain more accurate measurements. The shutter

Table 1 ARFF file organization. The coefficients of features along with the other parameters are arranged along the columns. In contrast, each pixel of the selected RoI is arranged along the rows.

	Transform 1 c_1, \dots, c_n	Transform N(optional) c_1, \dots, c_n	Position	Angle	Active brightness	Shutter	Output
$(u, v)_1$	*	*	*	*	*	*	*
\vdots							
$(u, v)_k$	*	*	*	*	*	*	*
\vdots							
$(u, v)_N$	*	*	*	*	*	*	*



Fig. 2 The materials employed in our tests are in the order: iron, fir wood, plywood, plastic, polystyrene, reflective surface, white fabric, dark fabric, aluminum, and glass.

value is the time duration in which the camera chip is exposed to IR light. This value is expressed in milliseconds (ms).

The last column represents the expected output value, which is identified by a cardinal number. In this regard, a specific number is assigned to represent each material to be classified (see Fig. 2 for more clarity).

Finally, it is worth noting that each row of the file refers to the single pixel belonging to the extracted RoI. Furthermore, as shown in Table 1, more than one feature transform can be employed. In this regard, some tests will show how the classification is affected.

It is important to underline that the data given from the acquisitions have been split into two categories: the training and the validation set. Specifically, 20% of the entire dataset has been reserved to represent the validation set. In this way, the validity of the method has been tested by analyzing the predictive accuracies of recognition.

4 Experimental Results

This section will explain the obtained outcomes by considering the analysis of 10 different materials (refer to Fig. 2). Three sets of acquisitions have been separately performed:

- analysis of eight materials by considering fixed panel poses (see Sec. 4.1);
- analysis of four materials by changing both the orientation and the displacement of the panel with respect to the ToF sensor (see Sec. 4.2);
- analysis of eight materials by introducing another wooden panel in the dataset (see Sec. 4.3).

All data belonging to one class have been collected from one panel. Specifically, only the test described in Sec. 4.3 employs a dataset where two different types of wood are used to represent such a category.

Moreover, only planar surfaces have been analyzed to prove the validity of method. Nonsolid materials, such as the dark and white fabric, have been fastened onto flat panels. Therefore, all reported outcomes refer only to planar materials. However, the presented methodology could be extended to the analysis of nonregular surfaces as well. In this regard, different patches having planar surface geometry can be detected from an object under investigation. Exploiting the definition of the shape index (SI) as in Ref. 15, it is

possible to measure the surface shape of any point belonging to a patch of interest. In detail, convex surfaces are identified by large SI values, concave surfaces have small SI values, and planar surfaces have medium SI values. Consequently, the patches having a medium SI value can be employed by our method to classify the material of items.

For the sake of completeness, some relevant optical properties related to examining materials are reported in Table 2. In this regard, the more common materials having different reflectance and refractive indices has been investigated to evaluate our approach.

The experimental setup is reported in Fig. 3. At this stage, a brief discussion about ambient light conditions is due. In general, ToF sensors suffer from issues linked to background

Table 2 Main optical characteristics of employed materials obtained by considering the wavelength of our ToF sensor (Fotonic E70 having $\lambda = 850$ nm). The database available online³¹ has been used to derive such parameters.

Material	Refractive index	Extinction coefficient	Reflectance
Iron	2.9541	3.4658	0.5712
Wood ^a	1.4680	—	0.0359
Plastic	1.5248	0.0018	0.0432
Polystyrene	1.5867	—	0.0514
Reflective surface	1.5162	—	0.6277
Fabric (cotton) ^b	1.5346	—	0.4501
Aluminum	2.5112	8.0136	0.8015
Glass	1.5162	3.1×10^{-6}	0.0350

^aIn the following experiments, there are two kinds of wood, i.e., fir wood and plywood, which have different properties due to different fabrication methodologies and thicknesses. Here, the properties of fir wood are reported.

^bTwo different-colored fabrics have been considered: dark and white. The optical properties will change since the absorption will depend on the material pigments. The physical properties reported in the table ignore the material color.

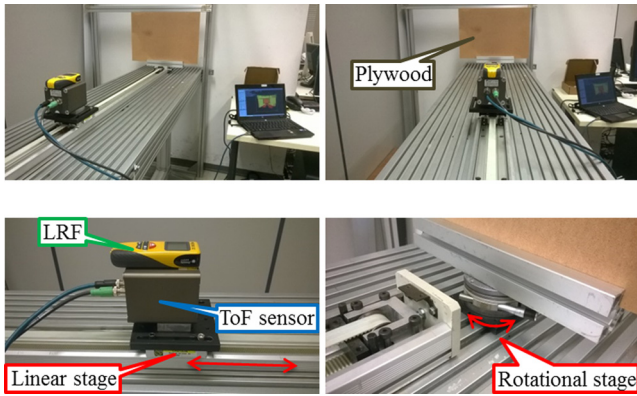


Fig. 3 Experimental setup used in our tests. The acquisitions have been performed without external and artificial lights. (a) Side view and (b) rear view of acquisition system. (c) Enlargement of translational stage along with ToF camera and laser range finder. (d) Enlargement of rotational stage used to provide rotational movements.

light. In fact, external light sources, such as sunlight or artificial lighting, could produce significant degradation of measurements. Therefore, outer optical band-pass filters could be adopted to selectively transmit only the rays having the expected wavelength, which will be input to the range camera. Hence, meaningless light is physically filtered out of the computation. In this regard, the employed ToF camera already mounts an internal optical filter. Nevertheless, to consistently reduce errors due to ambient light conditions, the experiments have been run in a completely controlled environment (without external and artificial lights).

The ToF camera used in our tests is the Fotonics E70,³² with a resolution of 160×120 pixels and a maximum measurement range of 7 m. The illumination unit emits modulated waves of near-infrared light (NIR), which are triggered by an internal reference signal, i.e., a sinusoid having a modulation frequency of 15 MHz and a wavelength of 850 nm. The phase of the incoming wave-front is estimated by means of the four-buckets algorithm, and the related distance is computed once the phase measurement is known.

It is worth highlighting that a bounded operating range of the ToF sensor has been employed to obtain more accurate measurements. In this regard, preliminary experiments not reported in this paper have shown that an average absolute distance error of 0.012 m is obtained by considering the

distance range of 0.5 to 3.5 m. Conversely, the error increases significantly when higher distances are considered. Hence, the presented tests have been performed taking into account only distances in this range.

The range sensor has been fastened onto a translational stage and then linked to a laptop. The target panel, i.e., the material under investigation, has been placed in front of the camera. Moreover, a rotational stage mounted under the panel is responsible for measuring the tilts as well as providing the rotational movements. The actual distance between the camera and the material is given by means of a dot laser range finder (LRF)³³ having an operating range of 0.1 to 10 m and a precision of 1 mm.

4.1 Eight Materials with Fixed Panel Poses

In this experiment, all materials except glass and fir wood have been considered. The position and the orientation of the panel on which the materials are fastened have been held fixed to 1 m and 0 deg, respectively. Furthermore, a sequence of 300 frames has been acquired.

Figure 4 shows the predictive accuracies given from the analysis of the eight materials. These accuracy rates have been obtained by considering the confusion matrices returned by Weka. Three different shutters have been tested. As shown by the bar graph, the shutter value heavily affects the likelihood of correctly recognizing the material type. In this regard, the DCT and HILBERT transforms seem less stable than the others. In contrast, the remaining transforms ensure higher recognition rates for all the reported shutter speeds.

The FOURIER, CHIRP, and KL transforms provide more reliable classification of the presented materials. Although the related recognition rates are comparable, the FOURIER transform requires less time to be computed with respect to the others. This aspect should be taken into account when real-time requirements for material recognition need to be addressed for specific applications.

As the bar graph shows, the KLT ensures high recognition rates. Such a transform was commonly employed in previous works^{34,35} for its effectiveness in feature extraction. As already stated, this transform is a representation of a linear combination of orthogonal functions such as the Fourier series, but the number of coefficients is variable. The Karhunen–Loève expansion is better suited, even though it requires more time to be computed with respect to the

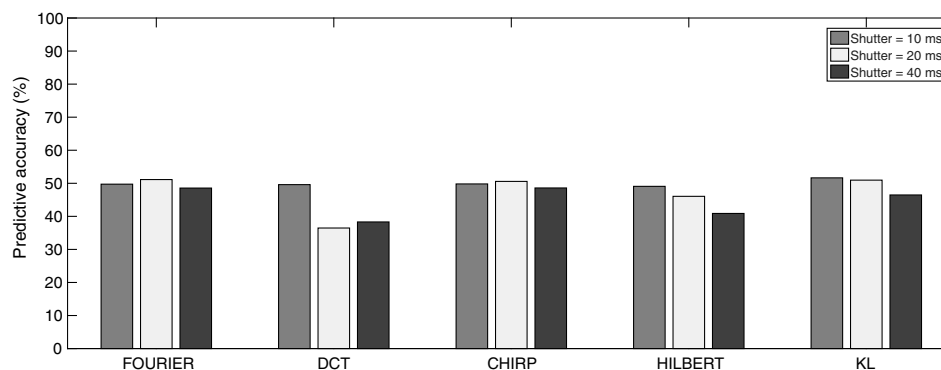


Fig. 4 Three different integration times have been evaluated for each employed feature. Note how the predictive accuracy consistently decreases for DCT and HILBERT transforms when an increase of shutter value occurs.

chirp and Fourier transforms. Therefore, the Fourier-based feature represents a good compromise between classification accuracy and computational time required.

In contrast, the DCT-based feature does not provide stable predictive accuracies for the considered shutter speeds. In this regard, the high-frequency noise that affects a signal is often isolated in a small number of coefficients. Since in our approach we take advantage of both the distance fluctuations and the intensity information for extracting a feature, it is fairly probable that such informative content is not preserved when the DCT transform is computed. Therefore, loss of information might occur when this transform is used.

Nevertheless, the classification rates obtained employing a unique feature do not compare favorably against previous methods. Therefore, experiments have been performed by providing the J48-based classifier with features extracted simultaneously by several transforms.

Hence, all the possible pair combinations of the presented transforms have been considered to enhance the predictive accuracies. Similar to the case of unique transform previously discussed, the features employing the DCT transform appear to be more affected by shutter value variations. As a matter of fact, a consistent decrease of predictive rates is observable from Fig. 5.

Conversely, the features extracted by using combinations of the FOURIER, CHIRP, and KL transforms ensure good predictive accuracies. Such results are still coherent with the ones obtained by employing a unique transform for extracting a feature of interest. In this regard, such features are less affected by the analyzed different shutter values.

Table 3 reports some important metrics useful in measuring the performance of the predictor. Among the computed features, only the most relevant has been proposed in this table. Moreover, a shutter of only 20 ms has been taken into account since that seems to ensure higher recognition rates.

As shown by the F -measure (i.e., the harmonic mean between precision and recall), the materials made of plastic and reflective surfaces can be predicted with good performance by means of KL-based features and CHIRP-KL-based features. In contrast, the dark fabric sample presents the worst case of this classification task due to its high absorption property. Other materials such as aluminum

and iron show average F -measure values since the considered shutter causes sensor saturation. Moreover, considering the average rank of the F -measure, the FOURIER-KLT combination and CHIRP-KLT show the best results. In general, if the KL score is low, most of the descriptors perform weakly.

Some of presented features enable recognition of the material type with good accuracy. However, the computational time required to calculate them has to be taken into account when dealing with real-time applications. For the sake of completeness, Table 4 reports the computational times needed to compute the features of interest.

The values in bold benchmark the elapsed time for computing the features that have shown the best results in terms of predictive accuracy. As observable, the FOURIER-based feature requires the least time to be computed with respect to the others. In fact, the FFT has been employed to obtain such features. In contrast, the features based on the KLT need more time to be extracted. In this regard, one order of magnitude higher than the FOURIER- and CHIRP-based features is evident from this table.

4.2 Four Materials with Different Panel Poses

In this case, a subset of considered materials has been examined since several panel poses have been taken into account, as shown in Table 5. This has been necessary for limiting the number of material/distance/orientation combinations to be acquired. Specifically, the plywood, the plastic panel, the reflective surface, and the glass surface have been analyzed since these materials are commonly present in indoor environments.

It is worth highlighting that the glass surface has been introduced in this experiment because it is more challenging for it to be acquired by the ToF camera, like the reflective surface. Therefore, its analysis could be of more interest than those of other material types.

Table 5 reports the different attribute values used during the experiment. In other words, several acquisitions have been performed considering the combinations of values listed in the table.

It is worth noting that the position and shutter values are related. In other words, when the panel is placed close to the ToF sensor, small shutter values are needed to get reliable

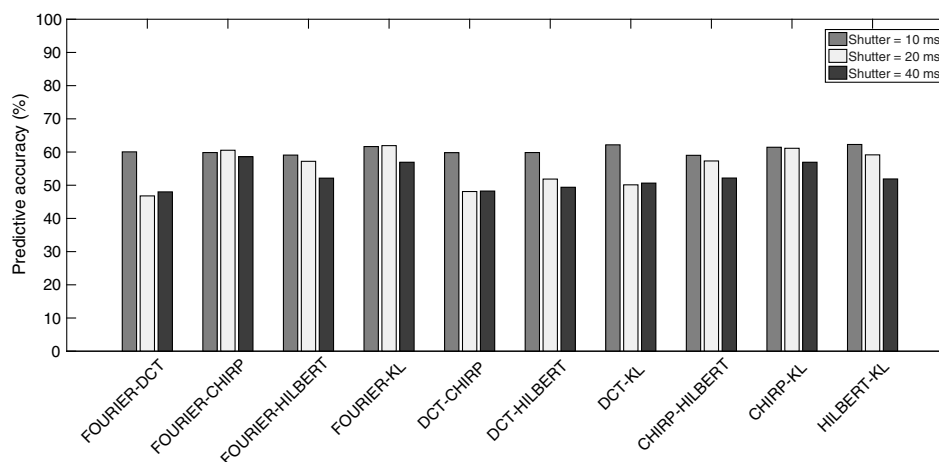


Fig. 5 The bar graph shows how the combination of transforms allows an increase of recognition rates of materials despite shutter changes.

Table 3 The true positive (TP), true negative (TN), false positive (FP), false negative (FN), precision, recall, and F -measures are reported for each material and extracted feature combination.

Material	Feature	TP	TN	FP	FN	Precision (%)	Recall (%)	F -measure (%)	Rank F -measure
Aluminum	FOURIER	110	2521	124	124	47.01	47.01	47.01	3
	CHIRP	111	2514	131	123	45.87	47.43	46.62	4
	KL	107	2499	146	127	42.29	45.72	43.93	5
	FOURIER-CHIRP	146	2515	130	124	52.89	54.07	53.44	2
	FOURIER-KL	169	2487	158	101	51.68	62.59	56.65	0
	CHIRP-KL	169	2482	163	101	59.00	62.59	56.14	1
Iron	FOURIER	114	2403	203	159	35.96	41.75	38.64	4
	CHIRP	120	2400	206	153	36.81	43.95	40.07	3
	KL	100	2412	194	173	34.01	36.63	35.27	5
	FOURIER-CHIRP	156	2396	210	153	42.62	50.48	46.22	0
	FOURIER-KL	126	2440	166	183	43.15	40.77	41.93	2
	CHIRP-KL	127	2437	169	182	42.91	41.12	41.98	1
Wood	FOURIER	147	2374	197	161	42.73	47.72	45.09	4
	CHIRP	132	2384	187	176	41.38	42.86	42.10	5
	KL	153	2400	171	155	47.22	49.67	48.42	2
	FOURIER-CHIRP	168	2384	187	176	47.34	48.83	48.07	3
	FOURIER-KL	190	2378	193	154	49.60	55.23	52.27	0
	CHIRP-KL	185	2376	195	159	48.68	53.77	51.17	1
Plastic	FOURIER	268	2465	54	92	83.22	74.44	78.59	4
	CHIRP	268	2465	54	92	83.22	74.40	78.59	4
	KL	297	2482	37	63	88.92	82.51	85.59	0
	FOURIER-CHIRP	304	2465	54	92	84.92	76.77	80.64	3
	FOURIER-KL	330	2466	53	66	86.16	83.33	84.72	1
	CHIRP-KL	330	2466	53	66	86.16	83.33	84.72	1
Polystyrene	FOURIER	241	2081	273	284	46.88	45.79	46.39	4
	CHIRP	246	2081	273	279	47.39	46.85	47.13	3
	KL	233	2056	298	292	43.88	44.38	44.13	5
	FOURIER-CHIRP	283	2079	275	278	57.20	50.45	50.59	1
	FOURIER-KL	288	2081	273	273	51.34	51.34	51.37	0
	CHIRP-KL	278	2078	276	283	51.80	49.55	49.87	2

Table 3 (Continued).

Material	Feature	TP	TN	FP	FN	Precision (%)	Recall (%)	F-measure (%)	Rank F-measure
Reflective surface	FOURIER	278	2558	21	22	92.98	92.67	92.82	3
	CHIRP	278	2558	21	22	92.98	92.67	92.82	3
	KL	281	2550	29	19	96.50	93.67	92.13	4
	FOURIER-CHIRP	314	2558	21	22	93.73	93.45	93.59	2
	FOURIER-KL	313	2567	12	23	96.30	93.15	94.71	1
	CHIRP-KL	313	2567	12	23	96.31	93.15	94.72	0
White fabric	FOURIER	233	2066	333	247	41.16	48.54	44.55	3
	CHIRP	217	2069	330	263	39.67	45.20	42.26	4
	KL	209	2050	349	271	37.46	43.54	40.27	5
	FOURIER-CHIRP	252	2070	329	264	43.37	48.84	45.94	1
	FOURIER-KL	254	2079	320	262	44.25	49.22	46.60	0
	CHIRP-KL	246	2064	335	270	42.34	47.67	44.85	2
Dark fabric	FOURIER	81	2278	202	318	28.62	20.31	23.75	5
	CHIRP	84	2259	221	315	27.54	21.05	23.86	4
	KL	87	2292	188	312	31.62	21.80	25.82	3
	FOURIER-CHIRP	120	2262	218	315	35.51	27.59	31.05	0
	FOURIER-KL	113	2271	209	322	35.09	25.98	29.85	1
	CHIRP-KL	112	2276	204	323	35.44	25.75	29.83	2

Table 4 Computational times obtained to compute the features of interest. All the transforms have been obtained using the software MATLAB. These outcomes refer to the analysis of one pixel belonging to plywood over time.

Feature name	Elapsed time (ms)
FOURIER	0.0382
DCT	0.2928
CHIRP	1.2265
HILBERT	0.2229
KL	10.5698
FOURIER-DCT	0.3224
FOURIER-CHIRP	1.2604

Table 4 (Continued).

Feature name	Elapsed time (ms)
FOURIER-HILBERT	0.2526
FOURIER-KL	10.5486
DCT-CHIRP	1.5175
DCT-HILBERT	0.4786
DCT-KL	10.7900
CHIRP-HILBERT	1.4746
CHIRP-KL	11.7152
HILBERT-KL	10.7169

Table 5 List of parameter values employed during our experiments.

Position values (m)	Rotation values (deg)	Shutter values (ms)
0.50	-45	3
0.75	-30	5
1.00	-10	10
1.25	0	15
1.50	10	20
1.75	30	40
2.00	45	—

depth measures. In contrast, the more the position increases its value, the more the shutter can be increased. Therefore, by considering a short distance between the panel and the acquisition system, small values of the shutter have to be chosen, thus enabling the avoidance of unwanted saturations of photosensors.

As shown in Fig. 6, an average increase of 10% in prediction accuracy has been obtained with respect to results reported in Fig. 4, since there are fewer material types that need to be considered. In this case, all features appear to be stable against shutter variations. The features based on FOURIER, CHIRP, and KL transforms show better predictive rates, although of a small margin.

By combining the transforms, an increase of predictive rates has been achieved, as shown in Fig. 7. Such outcomes prove once more that features extracted from multiple domains are better suited to be interpreted by the classifier.

This is consistent with the results obtained in the previous experiment. Moreover, a shutter value of 20 ms provides the highest predictive rates among those considered. In fact, this value seems to be the most suitable for the considered distance range and the materials used during the tests.

Table 6 reports the main metrics obtained by limiting the experiment to four material types only using a shutter value of 20 ms. As highlighted, very high *F*-measure values have been achieved by using the FOURIER-CHIRP-based features and FOURIER-KL-based features. Conversely, lower predictive rates have been obtained when features based on unique transforms have been used. In fact, as already stated, the combination of features enables enhancement of the predictive rates for material classification.

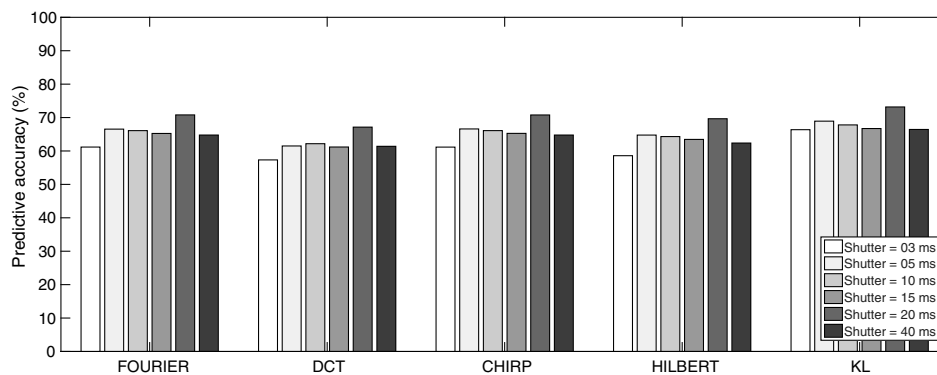
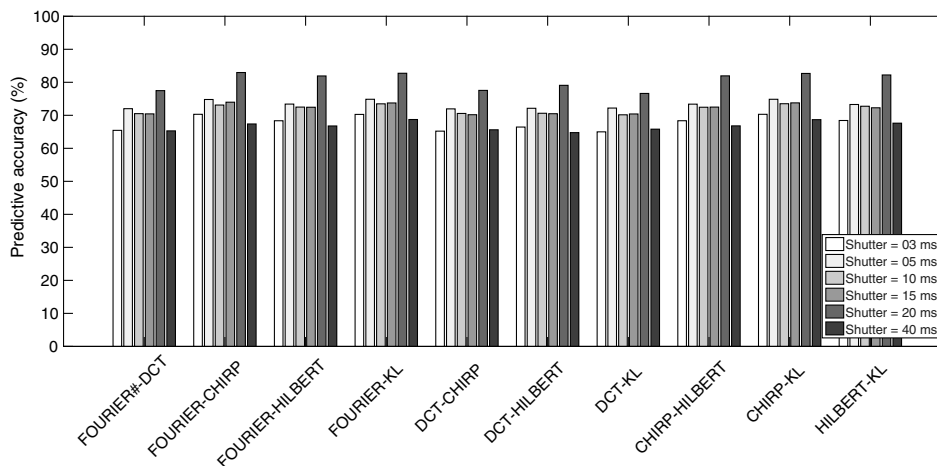
**Fig. 6** The bar graph reports the predictive accuracies by considering several poses of panel and different integration times. The highest recognition percentages are achieved with a shutter value of 20 ms.**Fig. 7** The combination of transforms involves an increase of predictive accuracies of about 10% with respect to those computed by considering a unique domain of transform.

Table 6 The main metrics for evaluating the classifier have been reported for the case of analysis. A shutter value of 20 ms has been considered to collect the data presented here.

Material	Feature	TP	TN	FP	FN	Precision (%)	Recall (%)	F-measure (%)	Rank F-measure
Wood	FOURIER	7297	4703	2136	2139	77.35	77.33	77.34	5
	CHIRP	7297	4705	2134	2139	77.37	77.33	77.35	4
	KL	7399	4858	1981	2037	78.88	78.41	78.64	3
	FOURIER-CHIRP	8211	5394	1445	1225	85.03	87.01	86.01	0
	FOURIER-KL	8176	5386	1453	1260	84.91	86.64	85.76	1
	CHIRP-KL	8173	5387	1452	1263	84.91	86.61	85.75	2
Plastic	FOURIER	1797	10,630	1930	1918	48.21	48.37	48.29	4
	CHIRP	1797	10,630	1930	1918	48.21	48.37	48.29	4
	KL	1957	10,778	1782	1758	52.34	52.67	52.50	3
	FOURIER-CHIRP	2432	11,533	1027	1283	70.30	65.46	67.80	0
	FOURIER-KL	2428	11,498	1062	1287	69.57	65.35	67.39	1
	CHIRP-KL	2429	11,495	1065	1286	69.51	65.38	67.38	2
Reflective surface	FOURIER	977	14,199	540	559	64.40	63.60	64.00	4
	CHIRP	976	14,195	544	560	64.21	63.54	63.87	5
	KL	1068	14,214	525	468	67.04	69.53	68.26	3
	FOURIER-CHIRP	1314	14,484	255	222	83.74	85.54	84.63	1
	FOURIER-KL	1317	14,486	253	219	83.88	85.74	84.80	0
	CHIRP-KL	1315	14,481	258	221	83.59	85.61	84.59	2
Glass	FOURIER	1451	14,540	147	137	90.80	91.37	91.08	4
	CHIRP	1449	14,539	148	139	90.73	91.24	90.98	5
	KL	1485	14,609	78	103	95.00	93.51	94.25	3
	FOURIER-CHIRP	1545	14,641	46	43	97.10	97.29	97.20	1
	FOURIER-KL	1545	14,646	41	43	97.41	97.29	97.35	0
	CHIRP-KL	1540	14,644	43	48	97.28	96.97	97.13	2

The results of the second experiment, therefore, confirm that the descriptors based on FOURIER-KL and CHIRP-KL provide the most effective material classification.

Table 6 shows that the predictive accuracies of materials made of plastic are considerably lower than others. Specifically, the reported data refer to a shutter value of 20 ms. In this regard, saturations of photosensitive receivers occur when the plastic panel has been acquired by considering this exposure time. Therefore, meaningless measurements have been returned by the ToF camera. In contrast, the other materials have not involved any photosensor

saturations, thus providing more reliable measurements. Additionally, the measurement fluctuations obtained by analyzing these materials appear to be more distinctive and meaningful by the classifier. As a consequence, they have been better classified, as shown in the table. Further experiments will be conducted to better investigate these material categories.

In Fig. 8, the recognition rates are shown, taking into account the position changes. As already highlighted, the features that ensure the highest scores are those based on a combination of FOURIER, CHIRP, and KL transforms.

The predictive accuracies related to short distances are slightly lower with respect to other positions, although by a small margin. This is probably due to the shutter values employed for these positions. Specifically, only integration times of 3, 5, 10, and 15 ms have been considered for such positions. However, it has been demonstrated that a shutter value of 20 ms ensures the highest recognition scores among those remaining. As a consequence, the scores of Fig. 8 are indirectly affected by the shutter value.

Similar to cases of analysis of shutter and position changes, the correlation between the angle variation and the predictive accuracy has been investigated. In this regard, Fig. 9 shows the obtained recognition rates by varying the angle value. It is important to underline that only the positive angles have been reported since the negative ones present very similar scores.

The most stable features against angle variations are still those obtained in previous experiments, namely the FOURIER–CHIRP, FOURIER–KL, and CHIRP–KL. High predictive accuracies have been achieved in this case as well.

In summary, it is possible to state that the combination of features enables enhancement of the prediction rates.

Moreover, three of these features can be considered more stable than others in material classification, even when different shutter values and panel poses are taken into account.

4.3 Eight Materials by Introducing Another Wooden Panel in the Dataset

The presented experiments have shown the efficiency of the proposed methodology for recognizing the material type. In this paper, the most common material categories have been considered. However, several objects belonging to the same category can be found when an environment is explored. For example, there are several surface typologies made of iron, wood, plastic, or fabric having different surface smoothness or reflective properties.

Nevertheless, it is unthinkable to classify all the existing materials. In this regard, a large dataset should be used to enhance the recognition rate of material. In fact, a wide set of objects made of the same material has to be considered during the training stage for each category.

In this experiment, another wooden panel made of fir, having different characteristics from the others (plywood),

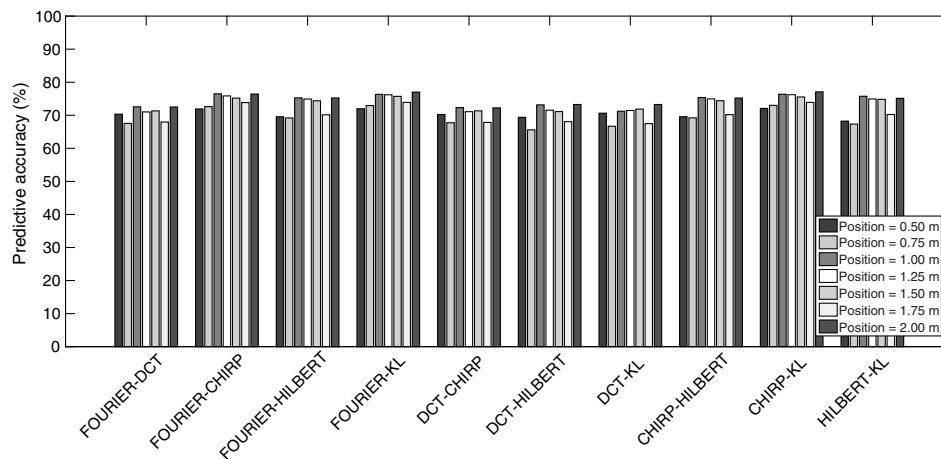


Fig. 8 The recognition rates of variations of distance from the ToF sensor and the examined panels are reported. The features that ensure the highest and most reliable scores are still those based on the combination of FOURIER, CHIRP, and KL transforms.

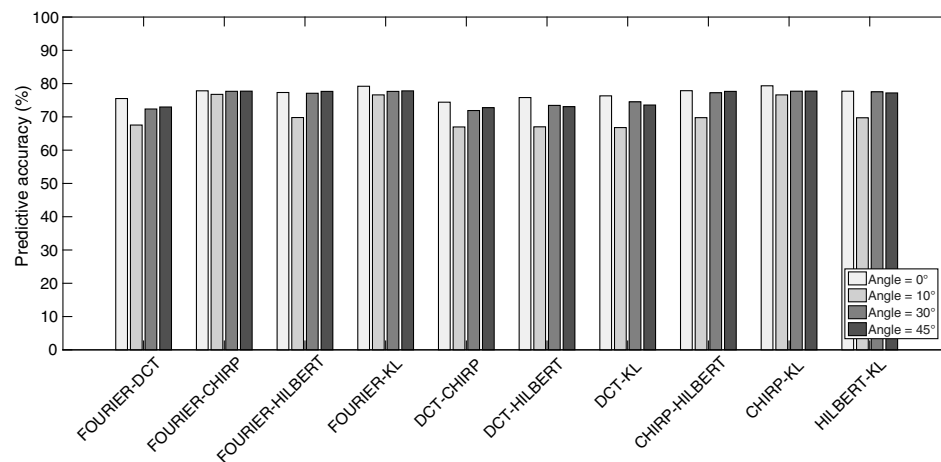


Fig. 9 Here, the predictive accuracies are shown by considering different orientations between the panel and the acquisition system.

has been analyzed. In this regard, the fir panel has a lighter color with respect to plywood and is also less smooth.

At first, data related to the fir panel have not been included in the training step. Therefore, the material category should be identified by the built classifier using only the data related to the other eight panels (similar to tests reported in Sec. 4.1). However, a low-predictive accuracy ($\sim 20\%$) has been reached. As a consequence, the fir panel has not always been identified as wood. This likely happens because the diverse roughness of the two wooden panels could affect category recognition. Conversely, by considering the data of the fir panel in the training stage and classifying the panel as wood, a significant increase of the recognition rate is achieved.

Figure 10 reports the predictive accuracies by considering the fir panel in the computation as well. As observable, a slight increase of scores has been obtained in comparison with the outcomes reported in Fig. 4. The improvement of rates is probably due to a better representation of the class “wood.” In other words, the more data used in the training step, the more the classifier is able to represent the real world.

The combination of features enables higher recognition rates (see Fig. 11). The features based on the combination

of FOURIER, CHIRP, and KL show more stable scores with respect to those that significantly decrease the predictive accuracies when the shutter changes its value. Hence, such outcomes prove once more that these features are sufficiently distinctive and reliable for achieving category classification.

The metrics related to this experiment are reported in Table 7. As observable from the table, the features computed from a combination of transforms ensure the highest classification rates. By comparing these results with those of Table 3, it is possible to note that similar prediction scores have been achieved. Nevertheless, the scores referring to the category of wood are considerably improved. Specifically, the F -measure of wood is increased by about 30% for all the considered features. This enhancement is due to a better representation of the class under investigation.

4.4 Discussion

Many works have been proposed in recent years for solving the problem of material recognition. Many of the proposed methodologies are essentially based on color and texture analysis of 2-D images. Very few works tackle this topic by exploiting 3-D information. In this regard, our method is able to classify the material typology by exploiting both

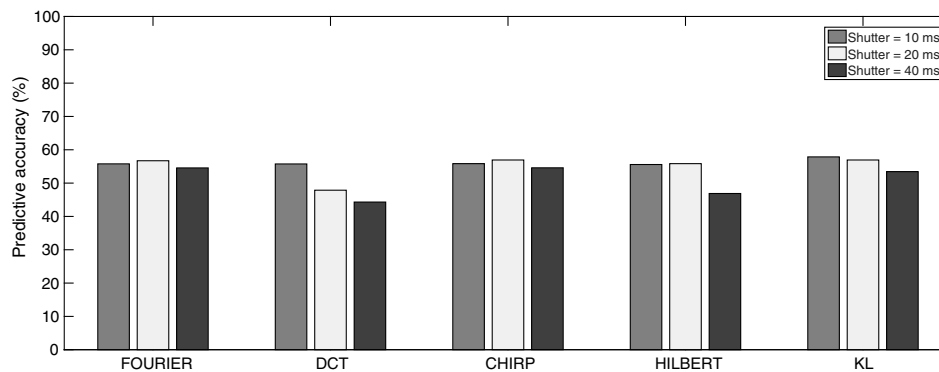


Fig. 10 Here, the recognition rates for each feature are reported by considering eight materials. Specifically, two typologies of wood (fir wood and plywood) have been taken into account to evaluate the accuracy of the classifier. The features based on FOURIER, CHIRP, and KL show more stable results against shutter value changes.

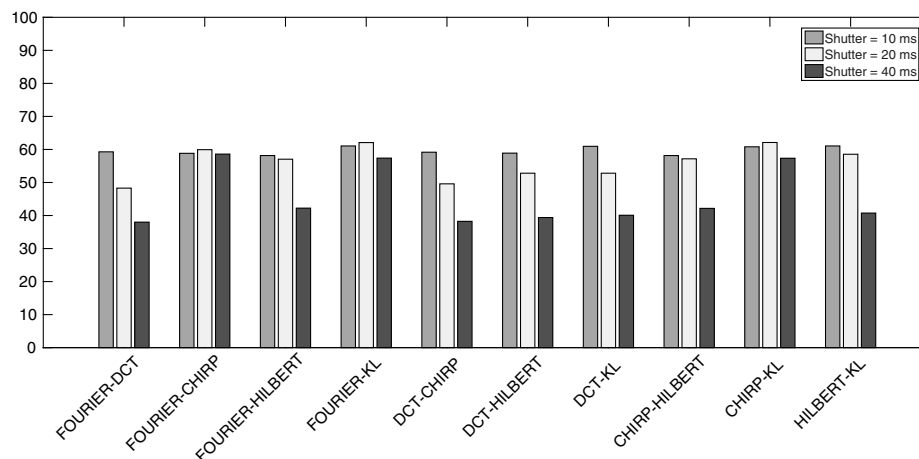


Fig. 11 The features based on FOURIER, CHIRP, and KL ensure the highest predictive accuracies. Moreover, these features are more reliable against shutter variations with respect to the remaining ones.

Table 7 Here, the main metrics related to this test are shown. Very high recognition scores for materials made of wood have been achieved in comparison with Table 3. Other predictive accuracies related to other categories are still comparable with those of that table.

Material	Feature	TP	TN	FP	FN	Precision (%)	Recall (%)	F-measure (%)	Rank F-measure
Aluminum	FOURIER	101	3178	127	120	44.30	45.70	44.99	4
	CHIRP	99	3180	125	122	44.20	44.80	44.49	5
	KL	113	3166	139	108	44.84	51.13	47.78	2
	FOURIER-CHIRP	112	3180	125	122	47.26	47.86	47.56	3
	FOURIER-KL	129	3168	137	105	48.50	55.13	51.60	0
	CHIRP-KL	129	3164	141	105	47.78	55.13	51.19	1
Iron	FOURIER	102	3052	214	158	32.28	39.23	35.42	5
	CHIRP	107	3056	210	153	33.75	41.15	37.09	3
	KL	96	3083	183	164	34.41	36.92	35.62	4
	FOURIER-CHIRP	120	3056	210	153	36.36	43.96	39.80	2
	FOURIER-KL	117	3100	166	156	41.34	42.86	42.09	1
	CHIRP-KL	119	3097	169	154	41.32	43.59	42.42	0
Wood	FOURIER	789	2384	187	166	80.84	82.62	81.72	1
	CHIRP	778	2377	194	177	80.04	81.47	80.75	5
	KL	804	2347	224	151	78.21	84.19	81.09	3
	FOURIER-CHIRP	791	2377	194	177	80.30	81.71	81.00	4
	FOURIER-KL	826	2347	224	142	78.67	85.33	81.86	0
	CHIRP-KL	808	2355	216	160	78.91	83.47	81.12	2
Plastic	FOURIER	255	3125	54	92	82.52	73.49	77.74	4
	CHIRP	255	3125	54	92	82.52	73.49	77.74	4
	KL	275	3138	41	72	87.03	79.25	82.96	2
	FOURIER-CHIRP	268	3125	54	92	83.23	74.44	78.59	3
	FOURIER-KL	295	3140	39	65	88.32	81.94	85.01	1
	CHIRP-KL	297	3140	39	63	88.39	82.50	85.34	0
Polystyrene	FOURIER	228	2723	291	284	43.93	44.53	44.23	5
	CHIRP	232	2729	285	280	44.87	45.31	45.09	3
	KL	227	2727	287	285	44.16	44.34	44.25	4
	FOURIER-CHIRP	245	2729	285	280	46.23	46.67	46.45	2
	FOURIER-KL	245	2782	232	280	51.36	46.67	48.90	1
	CHIRP-KL	251	2802	212	274	54.21	47.81	50.81	0

Table 7 (Continued).

Material	Feature	TP	TN	FP	FN	Precision (%)	Recall (%)	F-measure (%)	Rank F-measure
Reflective surface	FOURIER	266	3219	20	21	93.01	92.68	92.84	4
	CHIRP	266	3219	20	21	93.01	92.68	92.84	4
	KL	270	3220	19	17	93.43	94.08	93.75	2
	FOURIER-CHIRP	279	3219	20	21	93.31	93.00	93.16	3
	FOURIER-KL	282	3232	7	18	97.58	94.00	95.76	0
	CHIRP-KL	282	3232	8	18	97.24	94.00	95.59	1
White fabric	FOURIER	195	2737	322	272	37.72	41.76	39.63	4
	CHIRP	208	2744	315	259	39.77	44.54	42.02	3
	KL	169	2730	329	298	33.94	36.19	35.03	5
	FOURIER-CHIRP	221	2744	315	259	41.23	46.04	43.50	2
	FOURIER-KL	240	2721	338	240	41.52	50.00	45.37	0
	CHIRP-KL	241	2707	352	239	40.64	50.21	44.92	1
Dark fabric	FOURIER	73	2929	211	313	25.70	18.91	21.79	1
	CHIRP	72	2925	215	314	25.09	18.65	21.40	2
	KL	63	2944	196	323	24.32	16.32	19.53	4
	FOURIER-CHIRP	85	2925	215	314	28.33	21.30	24.32	0
	FOURIER-KL	63	2941	199	336	24.05	15.79	19.06	5
	CHIRP-KL	71	2935	205	328	25.72	17.79	21.04	3

the 3-D information and the intensity levels returned by the ToF camera.

Although the method proposed in this paper is directly comparable with just a few works, and, even then, it works on different datasets, a comparison table is shown in Table 8, with papers grouped according to the employed processing method.

Works in Refs. 1, 5, and 6 employed high-resolution 2-D cameras to acquire images with respect to the low resolution of our 3-D ToF sensor. In Ref. 6, an acquisition system based on a concave parabolic mirror and a beam splitter have been used to obtain the reflectance disks of materials. Different material typologies have been analyzed. In this regard, only Ref. 6 investigated similar materials. In contrast, Refs. 1 and 5 employed natural and building materials, respectively.

The work in Ref. 2 presents average scores even though several materials are examined. However, its methodology is applied to recognize natural materials, while this work is more concerned with structured indoor environments.

Finally, Refs. 14, 15, and 16 refer to the analysis of 3-D information for achieving material recognition. Among the discussed works, these are the closest to our approach

since they handled 3-D data to extract features. Furthermore, analogous material categories (wood, plastic, fabric, and paper) have been investigated. It is observable that a slightly lower accuracy rate has been achieved since in our case, challenging materials such as a reflective surface and glass have been considered.

Part of this advantage might also be dependent on the sensor's performance difference. In fact, the absolute accuracy (or the maximum systematic error on the distance measurements) and the repeatability (1σ) of SR4000³⁶ are equal to ± 10 and 4 mm, respectively. Conversely, the absolute accuracy and the repeatability (1σ) of Fotonic E70 are ± 20 and 7 mm. Hence, our acquisition system used to collect the data is less accurate than the other one.

5 Conclusions

In this paper, exploiting the information given by a ToF depth camera, a group of features has been computed to accomplish the task of material classification. For each material, a suitable RoI is considered. Specifically, all pixels belonging to this RoI are separately examined over time. A sequence of 300 frames is acquired for each material placed

Table 8 Overall recognition rates of discussed papers.

Category of method	Reference number	Type of features	Number of materials	Category of materials	Overall recognition rates (%)	
Color analysis and textural appearance	1	Local geometric and photometric properties	20	Natural	95.6	
	2	Texture	250	Natural	69.0	
	3	Gradient orientation	10	Flickr database	54.0	
	4	Color and curvature	10	Flickr database	53.1	
	5	Texture and statistical distribution filter response	20	Building	90.8	
	6	Reflectance with angular gradient computation	20	Real world	92.3	
	7 and 8	Texture	10	Flickr database	45.3	
	9 and 10	Multiscale 2-D	23	Real world	79.5	
	11	Intensity distribution	61	CUReT database	24.1	
	3-D data and intensity level analysis	14–16	Intensity and depth Information	4	Real world	86.7
		Our method	3-D data alterations and intensity Variable shutter and fixed panel pose	8	Real world	61.9 (Sec. 4.1)
		3-D data alterations and intensity Variable shutter and panel pose	4		82.9 (Sec. 4.2)	
		3-D data alterations and intensity Variable shutter and fixed panel pose	8		62.1 (Sec. 4.3)	

in front of the ToF sensor. At this stage, exploiting different transforms such as Fourier, Karhunen–Loève, chirp-z, and so on, different features have been extracted. Both a training and a validation dataset have been created in order to train and test a decision tree (J48) for classifying the materials.

Results have shown how the integration time (i.e., shutter value) affects the predictive accuracies of recognition in the event that only a unique transform domain is employed to classify materials. In this regard, the features based on the Fourier, chirp, and KL transforms seem more stable with respect to the shutter variations. By considering the combinations of transforms, a significant increase of recognition rates has been achieved as well. At the same time, by reducing the number of materials and introducing other information tied to the pose of the panel, predictive accuracies have slightly increased.

The efficiency of the presented methodology has been also proven by evaluating features with changes of the position and angle of panels. Good predictive rates have been achieved, confirming the stability of computed features against parameter variations.

Moreover, by enhancing the training dataset by introducing a new typology of panel in the category of wood (fir wood), significant recognition rates have been obtained, proving once more the effectiveness of our approach.

Since these results look promising, further work will be done to prove the robustness of the proposed methodology, e.g., by considering a wider set of materials and reducing the amount of frames acquired during an experiment. Moreover, other depth sensors such as Kinect *v1/v2* and Swiss Ranger 4000/4500 might be investigated in order to evaluate accuracy improvements. Finally, a method for patch extraction from objects will be considered for accomplishing material recognition in less controlled environments as well.

Acknowledgments

This work was funded within the CNR-ISSIA project “MASSIME-Mechatronic innovative safety systems (wired and wireless) for railway, aerospace and robotic applications.” The authors would like to thank Mr. Michele Attolico for technical support.

References

1. T. Leung and M. Jitendra, “Representing and recognizing the visual appearance of materials using three-dimensional textures,” *J. Comput. Vision* **43**(1), 29–44 (2001).
2. P. Vacha and M. Haindl, “Natural material recognition with illumination invariant textural features,” in *Pattern Recognition*, pp. 858–886, IEEE, Istanbul (2010).
3. D. Hu et al., “Toward robust material recognition for everyday objects,” in *British Machine Vision (BMVC)*, Vol. 13, pp. 1–11, BMVA Press, Dundee (2011).

4. I. Badami and K. Reinhard, "Material recognition: Bayesian inference or SVMs?," in *Seminar on Computer Graphics (CESCG)* (2012).
5. A. Dimitrov and G. F. Mani, "Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections," *J. Adv. Eng. Inf.* **28**, 37–49 (2014).
6. H. Zhang et al., "Reflectance hashing for material recognition," in *Computer Vision and Pattern Recognition (CVPR)*, pp. 3071–3080, IEEE, Boston (2015).
7. C. Liu et al., "Exploring features in a Bayesian framework for material recognition," in *Computer Vision and Pattern Recognition (CVPR)*, pp. 239–246, IEEE, San Francisco (2010).
8. L. Sharan et al., "Recognizing materials using perceptually inspired features," *J. Comput. Vision* **103**, 348–371 (2013).
9. S. Bell et al., "Material recognition in the wild with the materials in context database," in *Computer Vision and Pattern Recognition (CVPR)*, pp. 3479–3487, IEEE, Boston (2015).
10. S. Bell et al., "Material recognition in the wild with the materials in context database (supplemental material)," in *Computer Vision and Pattern Recognition (CVPR)*, IEEE, Boston (2015).
11. M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *Trans. Pattern Anal. Mach. Intell.* **31**(11), 2032–2047 (2009).
12. G. Sansoni et al., "State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation," *Sensors* **9**(1), 568–601 (2009).
13. D. Piatti, F. Remondino, and D. Stoppa, "State-of-the-art of TOF range-imaging sensors," in *TOF Range-Imaging Cameras*, pp. 1–9, Springer, Berlin Heidelberg (2013).
14. Md. A. Mannan et al., "Object material classification by surface reflection analysis with a time-of-flight range sensor," in *Advances in Visual Computing*, pp. 439–448, Springer, Berlin Heidelberg (2010).
15. Md. A. Mannan et al., "Material information acquisition using a ToF range sensor for interactive object recognition," in *Advances in Visual Computing*, pp. 116–125, Springer, Berlin Heidelberg (2011).
16. Md. A. Mannan et al., "3D free-form object material identification by surface reflection analysis with a time-of-flight range sensor," in *Machine Vision Applications (MVA)*, Nara, Japan, pp. 227–230 (2011).
17. S. Foix et al., "Lock-in time-of-flight (ToF) cameras: a survey," *Trans. Sensors J.* **11**(9), 1917–1926 (2011).
18. M. Lindner et al., "Time-of-flight sensor calibration for accurate range sensing," *Comput. Vision Image Understanding* **114**(12), 1318–1328 (2010).
19. S. Fuchs and G. Hirzinger, "Extrinsic and depth calibration of ToF cameras," in *Computer Vision and Pattern Recognition (CVPR)*, pp. 3777–3782, IEEE, Anchorage (2008).
20. M. Lindner and A. Kolb, "Lateral and depth calibration of PMD-distance sensors," in *Advances in Visual Computing*, pp. 524–533, Springer, Berlin Heidelberg (2006).
21. H. Rapp et al., "A theoretical and experimental investigation of the systematic errors and statistical uncertainties of time-of-flight-cameras," *J. Intel. Syst. Tech. Appl.* **5**(3/4), 402–413 (2008).
22. F. E. Nicodemus et al., *Geometrical Considerations and Nomenclature for Reflectance*, Vol. **160**, National Bureau of Standards, Washington, DC (1977).
23. H. W. Jensen et al., "A practical model for subsurface light transport," in *Computer Graphics and Interactive Techniques*, pp. 511–518, ACM, New York (2001).
24. S. Hui and S. H. Žak, "Discrete Fourier transform based pattern classifiers," *Bull. Pol. Acad. Sci.* **62**, 15–22 (2014).
25. N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *Trans. Comput.* **C-23**, 90–93 (1974).
26. K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press, San Diego (2014).
27. J.-Y. Bouguet, "Calibration Toolbox in Matlab," 2015, http://www.vision.caltech.edu/bouguetj/calib_doc/ (March 2016).
28. J. Heikkilä, "Geometric camera calibration using circular control points," *Trans. Pattern Anal. Mach. Intell.* **22**(10), 1066–1077 (2000).
29. Machine Learning Group at the University of Waikato, "Weka 3: data mining software in Java," 2002, <http://www.cs.waikato.ac.nz/ml/weka/> (March 2016).
30. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Francisco (1993).
31. M. Polyanskiy, "Refractive Index Database," 2008, <http://refractiveindex.info/> (March 2016).
32. Fotonic, "Fotonic e70 Data sheet," 2015, http://www.fotonic.com/wp-content/uploads/2015/10/Datasheet_Fotonic_E-series_2015_1-2.pdf (March 2016).
33. Prexiso, "Prexiso x2 data sheet," 2014, http://www.prexiso.com/en/prexiso-x2_54.htm (March 2016).
34. Y. Yamashita et al., "Relative Karhunen–Loeve transform method for pattern recognition," in *Pattern Recognition*, pp. 1031–1033, IEEE, Brisbane (1998).
35. J. Pannekamp et al., "Using the Karhunen–Loeve expansion for feature extraction on small sample sets," in *Industrial Electronics Society*, pp. 1582–1586, IEEE, Aachen (1998).
36. Mesa Imaging, "SR4000 Data Sheet," 2011, https://www.inf.u-szeged.hu/sites/default/files/ipcglab/docs/vision/SR4000_Data_Sheet.pdf (March 2016).

Fabio Martino obtained his bachelor's degree in electronic engineering from the Polytechnic University of Bari in 2010. At the same university, he attained a master's degree in electronic engineering in 2013 with a thesis titled "ToF range camera: analysis and evaluation." Since May 2013, he has been cooperating with ISSIA-CNR. His principal research interests are 3-D data analysis and image processing.

Cosimo Patruno received his BS and MS degrees in automation engineering from the Polytechnic University of Bari, Bari, Italy, in 2010 and 2013, respectively. Since May 2013, he has been cooperating with the Institute of Intelligent Systems for Automation (ISSIA), National Research Council of Italy (CNR), Bari, as a research collaborator. His main research interests include 3-D data analysis, signal and image processing, computer vision, and robotics.

Nicola Mosca received his degree (*cum laude*) in computer science from the University of Bari, Bari, Italy, in 2004, and his MPhil degree in transport systems engineering from the University of South Australia, Adelaide, Australia, in 2012. He has been cooperating with the Institute of Intelligent Systems for Automation (ISSIA) of the National Research Council since 2004. His research interests include image processing, computer vision, robotics, high-performance computing, and software design and development.

Ettore Stella received his degree (*cum laude*) in computer science from the University of Bari, Bari, Italy, in 1984. He is the scientific chief of several research projects. He is a coauthor of more than 100 papers in international journals and proceedings of conferences, book chapters, and international patents. From a professional point of view, he has certified experience in industrial automation, robotics, computer vision, high-performance computing, and software design and development.