

# MatTag: Practical material tagging using visual fingerprints

Adam Staš<sup>1</sup>, Daniel Pilar<sup>1</sup> and Jiří Filip<sup>1,\*</sup>

<sup>1</sup>The Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czech Republic

## Abstract

Assessment of material properties is essential for tasks such as similar material retrieval or swatch comparison in industrial design, manufacturing, and quality control. While many material similarity measures exist, they often fail to align with human perception. In this paper, we introduce a novel smartphone application using a machine learning model that leverages a perceptual representation known as the visual fingerprint of materials—linking image-based measurements to intuitive, human-understandable attributes. Trained on human ratings collected through psychophysical studies, the model can predict a material’s visual fingerprint using just two photographs captured under different lighting conditions. The application employs this model to assess any planar material sample using only a printed registration template and a flashlight. The app captures two photographs and predicts the material’s perceptual attributes. We demonstrate several practical use cases, including building personal material databases, retrieving visually similar materials, and exploring materials that match user-defined perceptual criteria. By enabling perceptually grounded comparisons and metadata extraction, our application provides a standardized representation of material appearance. This marks a step toward more intuitive and interoperable use of material properties across diverse digital environments.

## Keywords

Material Appearance, Perceptual Attributes, Visual Fingerprint, Smartphone Application, Material Retrieval

## 1. Introduction

There are currently many systems and data formats for the visualization and digital representation of real-world materials. However, these systems often lack meta-information about fundamental visual properties, which is essential for accurate material identification, comparison, and retrieval. Recent advances in the study of material perception have led to the development of a machine learning model that predicts how humans evaluate materials using a unified visual identifier a vector of sixteen perceptual attribute ratings. This perceptually grounded representation enables more intuitive interpretation of material appearance.

Despite this progress, there remains a lack of a practical tool that makes these technologies accessible to a broader audience – including researchers, domain experts, and advanced users in material analysis and visualization. Such a tool should support:

- Easy and consistent cataloging of personal/company material samples
- Searching for visually similar materials, either from a reference sample or via user-defined perceptual criteria
- Comparing the visual attributes of two selected materials in an intuitive way

These capabilities have significant potential in industrial applications, such as quality control across production batches, analyzing appearance differences between product lines, or selecting products that visually match customer expectations.

Mobile platforms are ideally suited for implementing such a tool due to their accessibility, integrated cameras, and ease of use—without the need for specialized hardware. In this paper, we introduce a mobile application for Android that builds on the aforementioned machine learning model to enable

MANER Conference Mainz/Darmstadt 2025 (MANER2025), August 29, 2025, Darmstadt, Germany

\*Corresponding author.

✉ filipj@utia.cas.cz (J. Filip)

🌐 <https://staff.utia.cas.cz/filip> (J. Filip)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

intuitive evaluation of real-world materials. The app features a user-friendly interface for capturing two photographs of a material sample under different lighting conditions. The resulting analysis is visualized as a polar plot, displaying predicted values of the sixteen perceptual attributes. Additional features include:

- Local storage of analyzed materials for offline access
- An interactive filtering interface based on the polar plot, allowing users to define desired perceptual properties by manipulating the attribute axes directly
- Cloud connectivity, enabling storage and search across a shared material database, with all analysis performed on a remote server

This application bridges the gap between advanced perceptual modeling and practical, accessible tools for material analysis in both professional and industrial settings. In the following sections we overview principle of material fingerprinting, its integration to a mobile application, evaluate its performance by capturing real-materials and outline several practical usecases.

## 2. Related work

Establishing a connection between perceptual texture spaces and computational representations has been the focus of numerous psychophysical studies, including work by Tamura et al.[1], Malik and Perona[2], Ravishankar and Jain [3], Mojsilovic et al.[4], and Long et al.[5].

More recent approaches combine human perception of materials—often represented through color and texture—with machine learning methods. Schwartz and Nishino [6] introduced the concept of visual material traits, which encode characteristic appearance features using convolutional representations of image patches for use in material recognition tasks. Subsequent work [7, 8] explored locally recognizable material attributes, learning perceptual attribute spaces from pairwise perceptual distances and training classifiers to reproduce them from image features.

To reduce reliance on predefined attribute sets, Schwartz and Nishino [9] later proposed a method for deriving material annotations by probing human visual perception through simple yes/no questions comparing pairs of small image patches.

Filip et al.[10] investigated perceptual dimensions of wood materials using video stimuli and a combination of similarity and attribute rating studies. In follow-up work[11], they linked these perceptual ratings with computational statistics and demonstrated the predictive power of statistical features on an independent test dataset.

While recent AI-based methods (e.g., CLIP [12]) provide fast multimodal embeddings for comparing visual content, they lack direct alignment with human perception of material appearance. For instance, although CLIP model itself can estimate the likelihood of certain visual attributes in provided image pairs, these estimates diverge significantly from human ratings. Yellow bars in Fig. 2 show correlations between model predictions and human ratings of the 16 visual attributes for different material categories. In average the correlation is only 0.321.

Deschaintre et al. [13] addressed this gap by linking free-text descriptions to fabric appearance. They collected a large dataset of human-written descriptions and corresponding fabric images, then derived a compact lexicon of attributes and structural descriptors used to describe fabric appearance. This annotated dataset was used to train a language-image model, producing a perceptually meaningful latent space for fabric retrieval and automatic annotation.

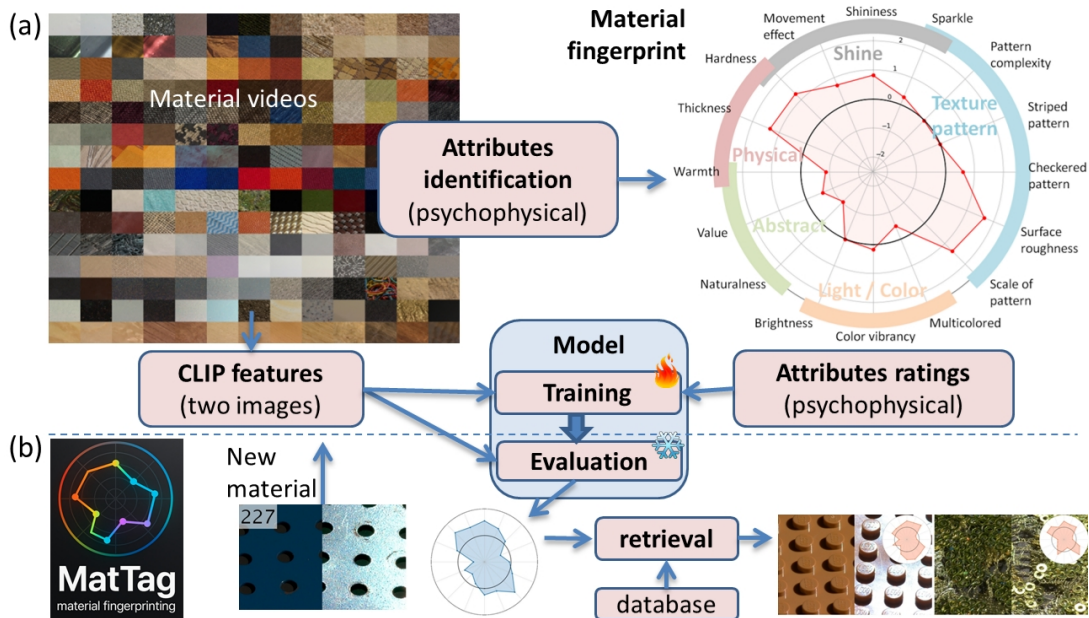
This paper is a follows up to a recent work by Filip et al. [14], which introduced a material fingerprinting model based on human-perceived visual attributes. The model predicts these attributes from only two photographs captured under controlled lighting conditions and forms the basis for the MatTag application introduced in this paper.

## 2.1. Material fingerprinting

This work builds on a previously introduced model for predicting key perceptual features of materials, as proposed in [14]. As outlined in Fig. 1-a, that study involved a psychophysical experiment using standardized video sequences of 347 diverse real-world materials, including fabrics, wood, and other surface types. The dataset was designed to cover a broad spectrum of textures, colors, and reflective properties.

Instead of relying on static imagery, the study used video sequences to capture the authentic appearance of flat material samples under varying viewing conditions. Sixteen key visual appearance attributes were identified through psychophysical analysis, supported by over 110,000 human ratings that mapped perceptual attributes across material categories.

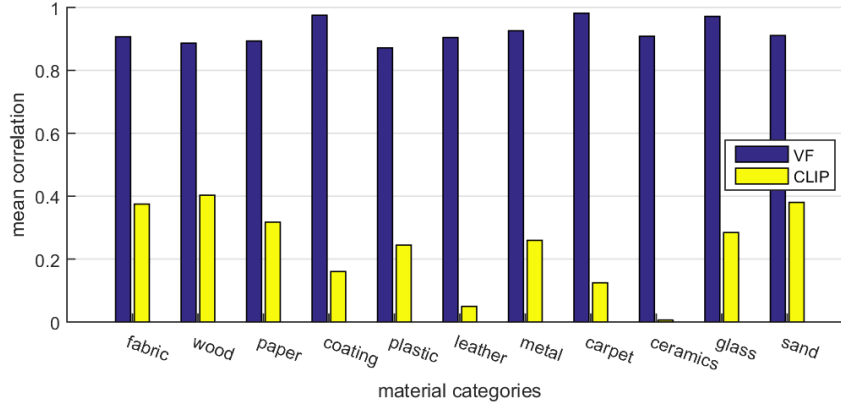
To model this perceptual space, CLIP-derived image features [12] were combined with a multi-layer perceptron (MLP) trained on the human rating dataset (see Fig.1-a). The resulting model predicts attribute ratings based on just two input photographs, taken under distinct lighting conditions: one emphasizing specular reflection, and the other capturing a non-specular view (Fig. 1-b). The predicted attribute values can be applied to tasks such as perceptually grounded material comparison or similarity-based retrieval.



**Figure 1:** Overview of the material fingerprinting model (a) and its evaluation workflow implemented in the MatTag application (b).

The predictive performance of the model is illustrated by the blue bars in Fig. 2, which show the correlation between model-predicted ratings and human ratings. Each bar represents the performance for a specific material category, while the overall mean correlation across all 68 test materials is 0.909. These test materials were not included in the model’s training set, ensuring a fair evaluation of generalization performance. It is worth noting that this level of performance is particularly impressive, given that the model predicts perceptual attributes from only two input images, whereas human observers based their ratings on full video sequences of the materials.

Evaluation of the model is performed server-side and typically completes within a few seconds. The method is robust to sample rotation, provided the material does not exhibit strong anisotropic effects. A current limitation lies in the training dataset, which, although carefully curated, is restricted to 347 materials. Nonetheless, this represents one of the largest sets used in perceptual material research to date. The model is relatively tolerant to minor deviations in capture geometry but may fail to distinguish between materials with appearance changes in angular configurations outside the sampled conditions. Furthermore, the trained model assumes a fixed capture scale and resolution, corresponding to a  $26 \times$



**Figure 2:** Comparison of two visual attribute prediction models with human ratings across several material categories. "VF" refers to our material fingerprint model [14], while "CLIP" uses CLIP-based embeddings [12]. Performance is measured as the mean correlation between ground truth and predicted 16-attribute vectors across all materials in each category.

26 mm material area resampled to  $512 \times 512$  pixels. As a result, it may not effectively evaluate materials with very low spatial frequency patterns or extremely gradual visual transitions.

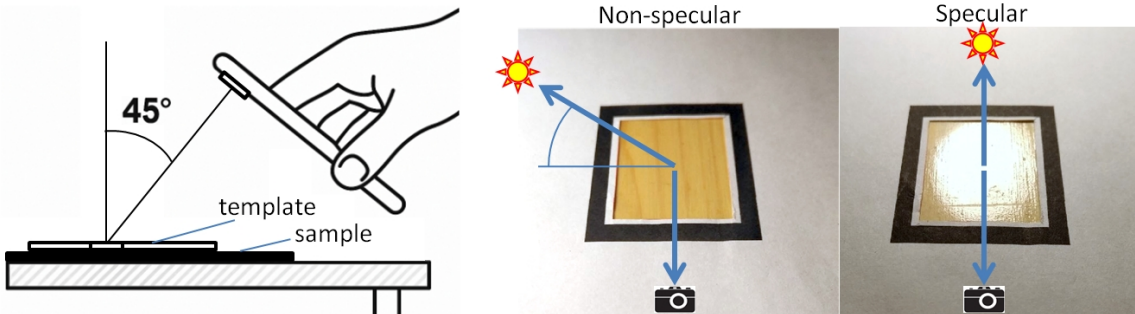
In this paper, we extend the use of the trained model by introducing a smartphone application, developed within master thesis project [15], that enables intuitive capture of the two required images and provides automatic evaluation of the material's perceptual fingerprint. We also describe several real-world use cases of the application.

### 3. MatTag application

Since the application predicts human judgments for a fixed set of visual attributes, it effectively performs attribute tagging. For this reason, we call it MatTag (Material Tag). The app consists of three main functional tabs: Capturing material images, Analyzing and comparing fingerprints, Application settings.

#### 3.1. Capturing workflow

The capturing process requires a printed paper template with a cut-out window to define the area of interest (see Fig. 3). This template is placed on the planar surface of the material being analyzed. It enables automatic detection of the correct camera polar angle, which is set to 45 degrees, matching the expected lighting geometry.

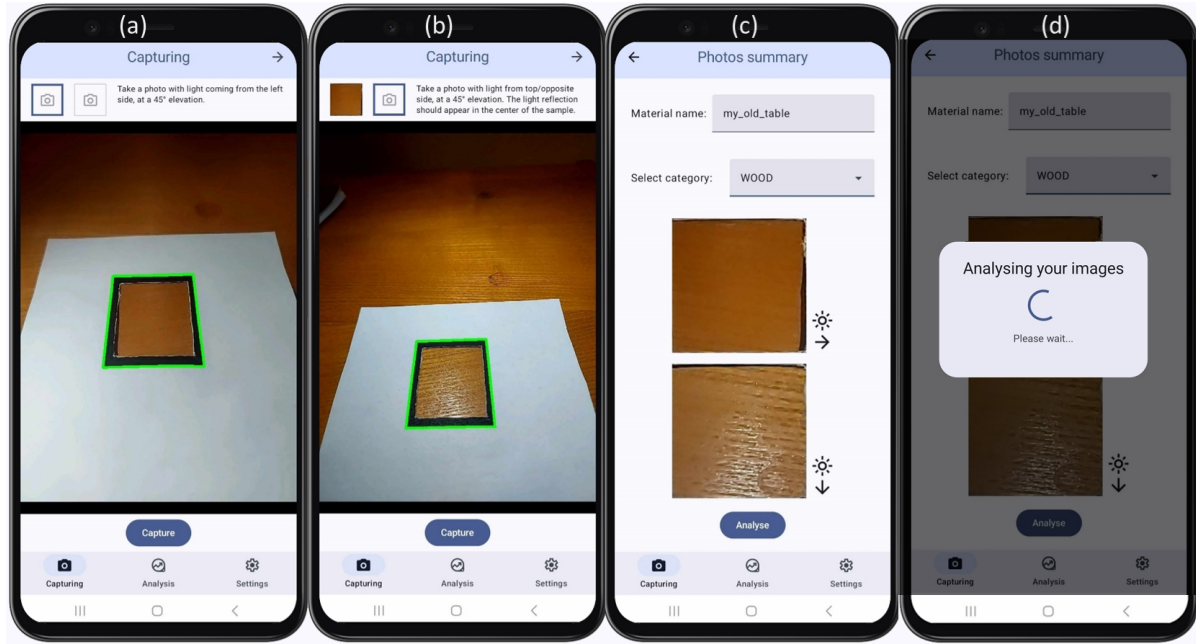


**Figure 3:** Material capture in the MatTag application using a smartphone, flashlight, and printed template.

Once the camera reaches the correct angle, the frame color changes from red to green, indicating that the user may proceed with capturing. Two images must be taken: One with the light coming from the left side, one with the light source placed opposite the camera (to emphasize specular reflection).



These steps are illustrated in Fig.4-a,b. After both images are captured, they are automatically corrected for perspective distortion and cropped. The user is then prompted to verify the image order (Fig.4-c) before initiating the analysis (d).



**Figure 4:** Screenshots of the material capture workflow (a,b) and image preprocessing steps (c,d).

### 3.2. Analysis and fingerprint comparison

During analysis, the images are uploaded to a remote server, where the visual fingerprint is computed in a matter of seconds. The resulting values for the sixteen perceptual attributes are returned and displayed as a polar plot (Fig. 5-a).

Once the fingerprint is available, users can perform:

- Retrieval of similar materials from either a local or a shared server-side database (b). The retrieval process uses the original distance function proposed in [14], which combines correlation and absolute differences between attribute values.
- Comparison of two materials, either in a single polar plot (c) or as two side-by-side plots (d)

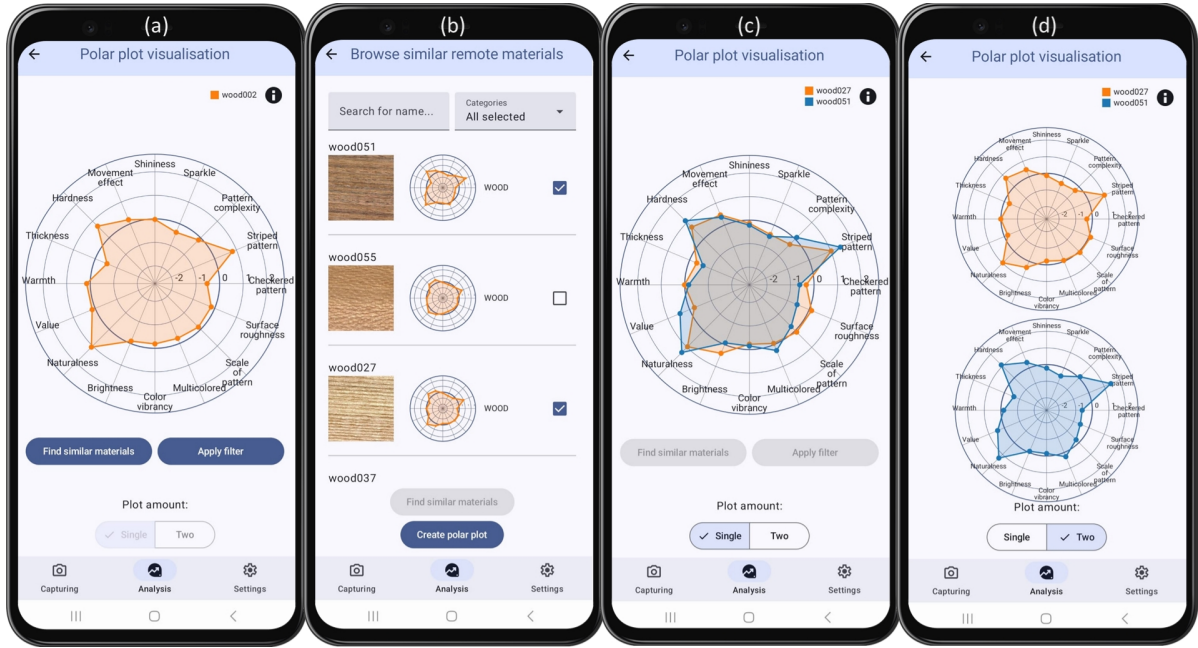
### 3.3. Interactive filtering of materials

Users can also filter materials by manually adjusting individual attribute values on the polar plot. This creates a custom query fingerprint used for perceptual retrieval. Each attribute can be increased or decreased by dragging its axis marker outward or inward, respectively, depending on how important the attribute should be in the search.

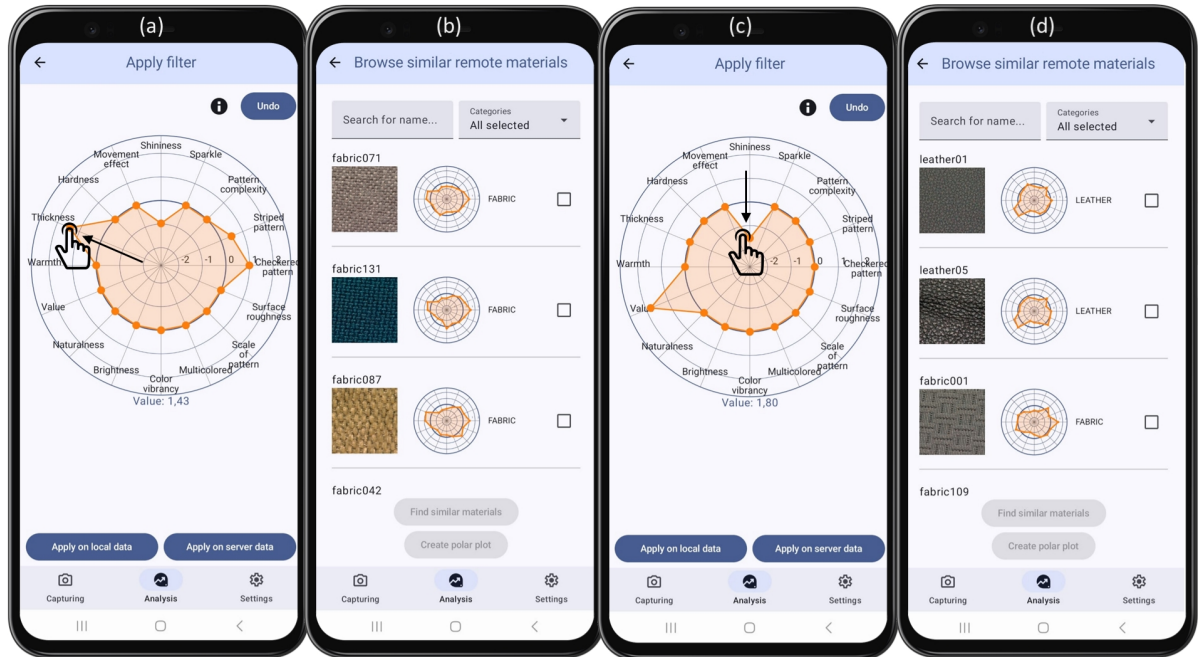
Fig. 6 illustrates retrieval results for two manually defined fingerprints. Examples (a,b) show retrieval for materials with higher thickness, striped or checkered patterns, and lower shininess. Examples (c,d) depict a search for materials with high brightness and low shininess. The last edited attribute is shown beneath the graph, and changes can be easily undone using the Undo button.

### 3.4. Platform and availability

MatTag is developed for the Android operating system and is compatible with devices running Android 12.0 (API level 31) or higher. The application architecture separates the user interface and capture



**Figure 5:** Visualization of material fingerprints (a), retrieval of similar materials (b), and side-by-side comparison (c,d).



**Figure 6:** Retrieval of materials (b,d) based on user-defined perceptual fingerprints (a,c).

logic from the perceptual analysis backend. The mobile front-end, developed in Kotlin and Java for Android, handles camera control, user interaction, and image preprocessing. The perceptual fingerprint is computed server-side using a Python-based backend, which runs the trained machine learning model and returns predicted attribute values to the app. Additional requirements are functional camera and internet connection for material analysis and remote storing materials (if approved). The app will be available through the project website, and we plan to release a stable version on the Google Play Store.

In the Settings tab, users can choose whether to allow their data to be stored on the remote server. If disabled, only local data is available for retrieval, while the server is used solely for computing the fingerprint. Additionally, users can export captured images and corresponding fingerprint values as a

CSV file for further analysis or archiving.

## 4. Results

The performance of the application was evaluated on a set of real-world material samples, as shown in Fig.7 and Fig.8. For each material, we present: the two input images (captured under specular and non-specular lighting), the estimated perceptual fingerprint visualized as a polar plot, and the five most similar materials retrieved from our database of 347 samples. Each displayed similar material includes both input images and the corresponding ground truth attributes obtained from human ratings. We also report the typicality of each material, visualized as a gray bar representing the mean absolute similarity score to the top 10% of most similar materials in the database. Although this typicality measure is not yet implemented in the current MatTag release, its integration is planned in a future update.

The results demonstrate the model’s rotation invariance, provided that the material’s main directional features are aligned with the template edges (see samples 8 and 9, 13 and 15, 14 and 16). The impact of lighting conditions, such as illumination intensity and direction, is also illustrated in samples 11 and 12, 13 and 14, and 15 and 16. It is important to note that polished wood materials are not retrieved, as the database does not currently contain such samples. Despite this, the method consistently retrieves perceptually relevant matches, even for materials that were not present in the training set. Examples include samples 2 (bubble foil), 5 (rough wood), 10 (cork panel), 21 (carbon fibers), and 22 (embossed rubber). Additionally, the method supports extended retrieval by allowing users to manually adjust specific visual attributes—e.g., to retrieve materials that are similar in most respects but with higher brightness or a stronger stripe pattern.

## 5. Practical recommendations and use-cases

**Illumination conditions** – The accuracy of the method depends on the consistent use of the same illumination source and its recommended positioning. Directional lighting is preferred over point-like sources, which are typical of many LED lights. Improved results can be achieved by placing a diffuse transparent material (e.g., frosted plastic or paper) a short distance in front of the LED to soften and expand the light.

For highly specular materials, it is recommended to slightly offset the light source from the exact specular direction. This avoids image saturation and improves visibility of surface texture, which is essential for accurate analysis.

A dark environment is not essential, but brightly lit rooms should be avoided. The intensity of the controlled light source should always be significantly higher than the ambient illumination.

**Sample placement** – To minimize blurring or distortion, the template must be positioned as close to the sample surface as possible and should remain perfectly flat without bending. While the method can still operate on images of nearly-flat surfaces, its accuracy may be reduced. In such cases, the approach is suitable only for relative comparisons between materials, provided they are captured using the same nearly-flat geometry.

Accurate alignment of the sample with the template, so that its edges are parallel to the camera’s optical axis, also improves attribute prediction, especially when comparing visually similar materials. This alignment is particularly important for anisotropic materials that exhibit directional structures, such as wood grains or woven fabrics. Note that the model cannot capture the full range of anisotropic behavior beyond what is represented in the two input images. To extend its capability to handle full anisotropy, one would need to broaden the angular trajectories of both lighting and viewing directions in the stimulus videos, repeat the rating experiment, and retrain the MLP model accordingly. Such an extension would also require a larger number of input images to sufficiently represent the material’s directional appearance.

Below, we outline several real-world use-cases for the MatTag application:

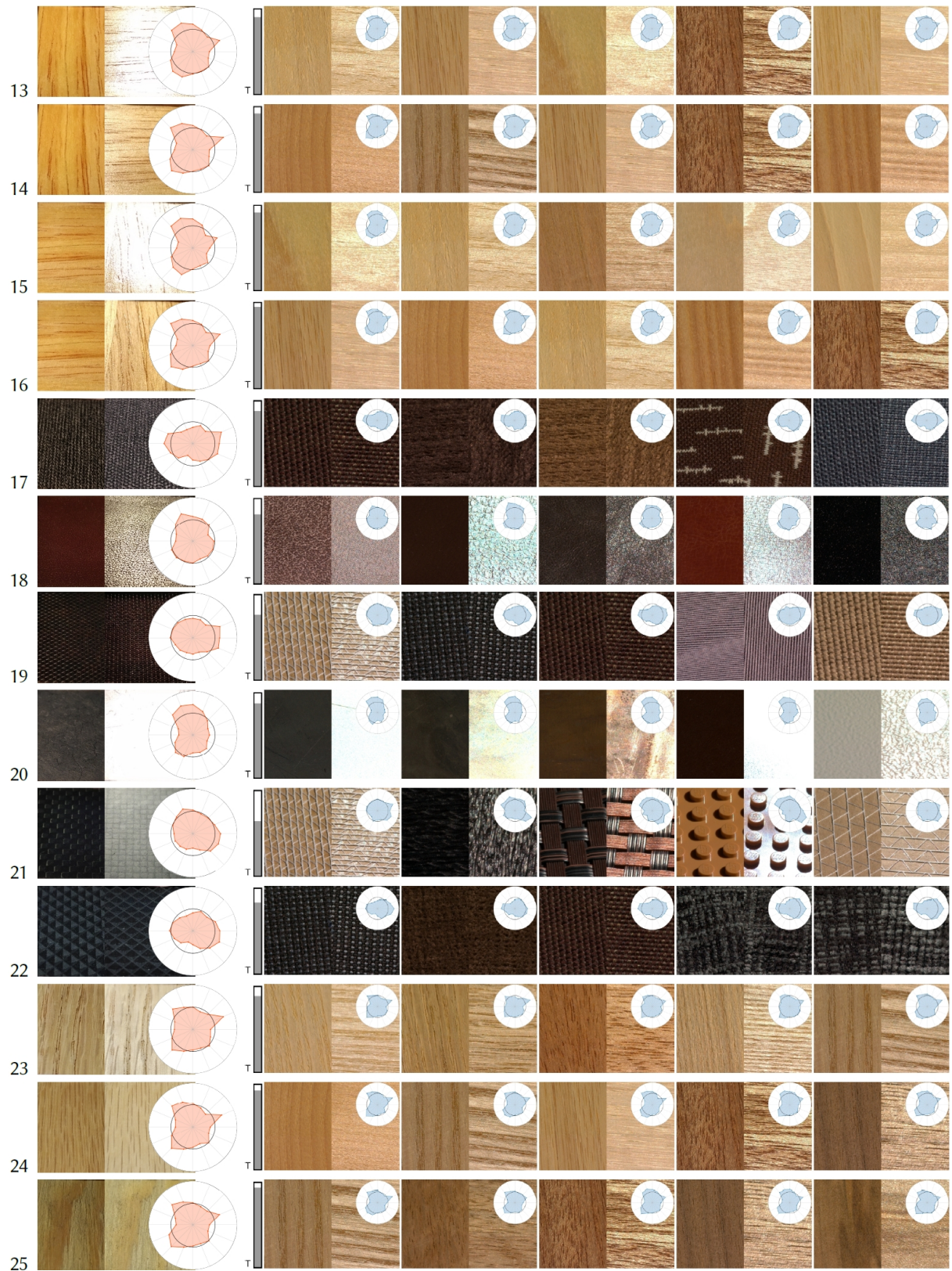




**Figure 7:** Material retrieval results: For each query material, two smartphone images captured using MatTag are shown (left) – one under non-specular and one under specular lighting. The five most similar materials retrieved from our database are displayed on the right. Material typicality is indicated by a gray bar.

1. **Creating a private material database:** When captured materials are stored locally, MatTag can be used to build a proprietary material database. Since the source code for the server-side component is publicly available, multiple users within an institution can share a private database of in-house materials for retrieval and similarity comparison. The reliability of such a database





**Figure 8:** Material retrieval results: For each query material, two smartphone images captured using MatTag are shown (left) – one under non-specular and one under specular lighting. The five most similar materials retrieved from our database are displayed on the right. Material typicality is indicated by a gray bar.



can be significantly improved by using consistent acquisition conditions – such as a smartphone holder and fixed lighting setup (e.g., two switchable lights at predefined positions). The computed visual attribute values can be exported and used as metadata, for example, to accompany digital material twins (BRDF, PBR, BTF) or to support structured material cataloging.

2. **Material retrieval:** Users can retrieve the most visually similar materials to a selected sample. The polar plot representation enables intuitive comparison of perceptual attributes, helping to explore subtle differences between materials stored locally or to evaluate the uniqueness of a sample when searching a remote database. In industrial settings, this can assist in identifying the closest match within a production portfolio to a customer’s reference material.
3. **Intuitive material filtering:** In addition to retrieval, users can interactively navigate material collections by adjusting the importance of individual perceptual attributes. This attribute-weighting approach supports exploratory browsing of both local and remote databases using an intuitive, perceptually grounded interface.

## 6. Conclusions

We presented MatTag, a practical mobile application for perceptual analysis and comparison of material appearance using their visual fingerprints. Built on a machine learning model trained on human perceptual data, MatTag predicts intuitive visual attributes from just two smartphone-captured images under controlled lighting conditions. The application enables users to capture, analyze, and retrieve materials based on their perceptual properties—without requiring specialized hardware. Through features such as interactive filtering, local and remote retrieval, and visual fingerprint comparison via polar plots, MatTag supports both professional workflows (e.g., material cataloging and quality control) and exploratory use in design and research. By offering a perceptually grounded, image-based representation of materials on a widely accessible platform, MatTag represents a step forward in bridging material perception research with practical, real-world applications in digital material workflows. In the future update we plan adding information on fingerprint prediction reliability.

## Acknowledgements

This research was supported by the Czech Science Foundation under grant GA22-17529S. We thank Veronika Vilímovská and Jan Kotera for their contributions to the implementation of the material fingerprinting model. We also gratefully acknowledge the volunteers who participated in the user study of the application.

## Declaration on Generative AI

During the preparation of this work, the author(s) used GPT-4 in order to grammar and spelling check. After using this service, the author(s) reviewed and edited the content as needed and take full responsibility for the publication’s content.

## References

- [1] H. Tamura, S. Mori, T. Yamawaki, Textural features corresponding to visual perception, *Systems, Man and Cybernetics, IEEE Transactions on* 8 (1978) 460–473. doi:10.1109/TSMC.1978.4309999.
- [2] J. Malik, P. Perona, Preattentive texture discrimination with early vision mechanisms, *JOSA A* 7 (1990) 923–932.
- [3] A. Ravishankar Rao, G. Lohse, Towards a texture naming system: Identifying relevant dimensions of texture, *Vision Research* 36 (1996) 1649–1669. doi:10.1016/0042-6989(95)00202-2.

- [4] A. Mojsilovic, J. Kovacevic, D. Kall, R. Safranek, S. Kicha Ganapathy, The vocabulary and grammar of color patterns, *Image Processing, IEEE Transactions on* 9 (2000) 417–431.
- [5] H. Long, W. Leow, A hybrid model for invariant and perceptual texture mapping, in: *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, IEEE, 2002, pp. 135–138.
- [6] G. Schwartz, K. Nishino, Visual material traits: Recognizing per-pixel material context, in: *2013 IEEE International Conference on Computer Vision Workshops*, 2013, pp. 883–890.
- [7] L. Sharan, C. Liu, R. Rosenholtz, E. H. Adelson, Recognizing materials using perceptually inspired features, *International Journal of Computer Vision* 103 (2013) 348–371.
- [8] G. Schwartz, K. Nishino, Automatically discovering local visual material attributes, in: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3565–3573.
- [9] G. Schwartz, K. Nishino, Recognizing material properties from images, *IEEE transactions on pattern analysis and machine intelligence* 42 (2019) 1981–1995.
- [10] J. Filip, J. Lukavský, F. Děchtěrenko, F. Schmidt, R. W. Fleming, Perceptual dimensions of wood materials, *Journal of Vision* 5 (2024) 12–12.
- [11] J. Filip, V. Vilimovská, Characterization of wood materials using perception-related image statistics, *Journal of Imaging Science and Technology* 5 (2023) 1–9. doi:10.2352/J.ImagingSci.Technol.2023.67.5.050408.
- [12] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, et al., Learning transferable visual models from natural language supervision, in: *International conference on machine learning*, PMLR, 2021, pp. 8748–8763.
- [13] V. Deschaintre, J. Guerrero-Viu, D. Gutierrez, T. Boubekur, B. Masia, The visual language of fabrics, *ACM Trans. Graph.* 42 (2023). doi:10.1145/3592391.
- [14] J. Filip, F. Dechterenko, F. Schmidt, J. Lukavsky, V. Vilimovska, J. Kotera, R. W. Fleming, Material fingerprinting: Identifying and predicting perceptual attributes of material appearance, *arXiv*, to appear in *Royal Society Open Science* (2024). doi:<https://arxiv.org/abs/2410.13615>.
- [15] A. Stas, Mobile application for evaluation of material visual properties, Diploma thesis, Czech Technical University in Prague, Prague, Czech Republic, 2025. URL: <https://dspace.cvut.cz/handle/10467/122804>, accessed: 2025-07-17.

## Online Resources

The source codes for the MatTag application and server are available:

- GitHub - MatTag application <https://github.com/adamstas/material-fingerprint-app>
- GitHub - MatTag server <https://github.com/adamstas/material-fingerprint-app-server>