Expressive-WICED | Artworks, Posters, and Demos
D. Berio and A. Bruckert (Editors)

EXPRESSIVE 2025

Robotic Painting using Semantic Image Abstraction

Michael Stroh[1], Patrick Paetzold[1], Daniel Berio[2], Frederic Fol Leymarie[2], Rebecca Kehlbeck[1], Oliver Deussen[1]
[1]University of Konstanz, Germany
[2]Goldsmiths University of London, United Kingdom

(a) (b) (c) (d) (e)

Figure 1: *Semantically guided image abstraction: a) Original image; b) semantic saliency map; c) color-coded semantic regions; d) resulting enhanced abstraction with varying degrees of abstraction (facades, ground, windows on tower); e) robot painting*

Abstract

We present a novel image segmentation and abstraction pipeline tailored to robot painting applications. We address the unique challenges of realizing digital abstractions as physical artistic renderings. Our approach generates adaptive, semantics-based abstractions that balance aesthetic appeal, structural coherence, and practical constraints inherent to robotic systems. By integrating panoptic segmentation with color-based over-segmentation, we partition images into meaningful regions corresponding to semantic objects while providing customizable abstraction levels we optimize for robotic realization. We employ saliency maps and color difference metrics to support automatic parameter selection to guide a merging process that detects and preserves critical object boundaries while simplifying less salient areas. Graph-based community detection further refines the abstraction by grouping regions based on local connectivity and semantic coherence. These abstractions enable robotic systems to create paintings on real canvases with a controlled level of detail and abstraction.

CCS Concepts

• ***Computing methodologies*** → ***Non-photorealistic rendering;*** *Image processing;*

1. Introduction

Non-photorealistic rendering (NPR) [KCWI13] encompasses techniques that transform scenes or images into abstract, artistic renditions. As a subset of image abstraction, NPR simplifies complex visual information into aesthetically pleasing and recognizable representations. One particular aspect focuses on creating painterly renderings, which can even be realized as physical artworks that emulate paintings created by human artists. Robotic painting systems extend the original digital NPR techniques to the physical domain by realizing such abstracted renderings as tangible artworks. By mimicking various painting styles, these systems provide insights into artistic processes replicating human-like conceptualization and craftsmanship, which are often only poorly (or not all) documented, or left to be intuitively reproduced by human artists [GD22].

Many existing NPR and robotic painting systems have been replicating painterly styles consisting of many visible brush strokes, such as impressionism and expressionism [LPD13, SMO23, TL12, TF13]. A common approach involves iteratively overlaying semi-transparent brush strokes of varying sizes. A visual feedback system determines if the local target color with re-

DOI: 10.2312/exw.20251070

spect to the original pixel color is achieved [LPD13, SMO23]. This can give acceptable results; however, it is quite limited in abstraction capabilities and flexibility, as the system mainly depends on the number of iterations and the selected brush types and sizes.

Rather than focus on the distribution of a sequence of brush strokes, we are concerned with region-based abstraction methods and show how our approach facilitates the generation and robotic reproduction of painterly styles with a strong underlying structured abstraction. These styles, characterized by large, uniformly filled areas and geometric forms, draw inspiration from the aesthetics of Pop Art and Orphism.

This shift toward region-based robotic painting introduces specific limitations for image abstractions intended for physical realization. (i) The regions should not overlap to avoid too many unplanned color interactions and reduce the risk of overpainting relevant features. (ii) The number of regions should not be too large due to physical time constraints as the color paints dry out. Also, the painting process should not require significantly more time than it would for a human artist. (iii) The regions should not be too small or have too complex geometry, as thin regions or small details might not be realizable due to brush-size limitations and unpredictable dynamics of the brush when in contact with the canvas during robotic movement. (iv) The semantic integrity of objects is critical as the robot fills with paint the regions; any fragmentation on the segmentation part could lead to rendering important objects unrecognizable with physical inaccuracies during the painting.

This paper presents a method to generate semantic image abstractions designed specifically for robotic painting. Unlike most digital abstraction techniques, our approach considers physical constraints in the design of the abstraction process. By leveraging semantic saliency maps and region-based segmentation, our method balances detail and abstraction to produce coherent, adjacent regions optimized for robotic painting. Figure 1 illustrates the core components and results of our method, starting from a natural scene (Figure 1a) which is progressively abstracted into a representation suitable for physical realization (Figure 1d). Figure 1e shows a painting created using an industrial robot based on our representation. Towards that, we propose a novel image abstraction pipeline that transforms pixel-based input images into vectorized abstractions tailored to robotic systems. By integrating semantic saliency and panoptic segmentation with adaptive region merging, our method addresses the practical challenges of robotic painting while reducing the reliance on manual parameter tuning. We demonstrate how these abstractions enable robotic systems to produce artistic paintings on physical canvases.

2. Related Work

Various image segmentation methods follow a two-step approach consisting of an initial over-segmentation of an input image, followed by a similarity merging of adjacent segments. They mainly differ in how they initially over-segmentate and how they define similarity between regions. Felzenszwalb and Huttenlocher [FH04] over-segment an image and iteratively merge similar neighboring segments until a desired level of abstraction is reached. While this approach is flexible regarding similarity measures, only color-based similarity has been used. Building on this idea of segment clustering, Achanta et al. [ASS*10] introduced *SLIC* (Simple Linear Iterative Clustering), which groups pixels based on both color and spatial proximity into compact and nearly uniform regions. Sadreazami et al. [SAM17] stylize images in a cartoon-like way by merging regions according to their color-based similarity in CIELAB color space combined with geometric region distance.

While such segmentation-based stylization methods effectively group image regions based on low-level features such as color, texture, or shape, they often lack semantic awareness or do not explicitly control the abstraction process in a way that aligns with robotic painting constraints. Many of these existing approaches prioritize perceptual similarity without ensuring structured, non-overlapping regions, leading to either excessive fragmentation, impractical for robotic painting, or the need for manual tuning to achieve meaningful abstraction.

Stroh et al. [SGD23] recently proposed a region-based image abstraction for robotic painting using panoptic segmentation employing a shape tree approach [FXDG17]. Their method greedily merges regions based on fixed color thresholds, followed by a heuristic filtering strategy that requires the labor-intensive selection of multiple user-defined parameters. Determining suitable abstraction levels for each panoptic object becomes particularly cumbersome in scenes containing many objects. This process demands significant manual effort, as thresholds must be adjusted individually for each object to achieve the desired level of abstraction.

3. Method

Our region-based image abstraction and stylization pipeline begins with an image over-segmentation, using the Felzenszwalb-Huttenlocher (FZH) graph-based color segmentation method [FH04], to create an initial set of minimally abstracted regions. This step efficiently produces regions that align with edges in the input image. Such edges are typically detected along local pixel color and luminance differences. However, this approach can be error-prone, for example, when bordering objects are separated by a soft brightness or color gradient, where the color changes gradually across large parts of the input image, and thus, no prominent border can be identified. In natural image scenes, the importance of image edges depends on the visual complexity of an area and the scale and number of features in the image context. This issue can be mitigated by considering image contents and semantics in the partitioning procedure. To this end, we combine the color-based over-segmentation with a semantic method that incorporates object-based scene understanding into our procedure. Specifically, we make use of the publicly available panoptic segmentation model OneFormer to identify relevant scene objects in the input image [JLC*23]. Figure 1c examplifies identified panoptic regions.

Traditional global abstraction approaches often fail to handle images with high variability in visual complexity. They are prone to merging semantically different regions or missing essential visual features. To address this, we propose a *locally adaptable image abstraction method* that balances low-level perceptual features, such as color similarity, with high-level semantic information. Our approach automatically selects suitable parameters to preserve critical details while meaningfully simplifying the image. Additionally, it provides the user with fine control over abstraction levels for different objects. At the core of our method lies the RAG, which represents regions and their relationships throughout the abstraction process. This graph is initialized with over-segmented regions computed by the FZH segmentation method [FH04], where each region corresponds to a node. Each such node stores the average color, derived from the region's pixels in the original image, and the semantic weight, computed by averaging pixel values from the saliency map.

We filter edges in the RAG based on their panoptic label and then weights to control the abstraction level and preserve key visual features. For each panoptic object, a threshold is applied to remove edges with large weights, effectively isolating regions with significant differences from their neighbors. This step prevents undesirable merges and ensures that object boundaries and critical transitions are maintained. For each panoptic subgraph, we compute the cumulative distribution function of edge weights. This function indicates the number of edges remaining in relation to the threshold values. Significant thresholds, corresponding to visible abstraction changes, are identified using *knee-point detection*. Knee-points indicate a substantial change from steep to gentle curvature in a function, which can also be interpreted as the transition between informative data points and noise [SAIR11]. These thresholds balance region merging with boundary preservation and can be refined by user-defined minimum and maximum abstraction values.

Even after edge filtering, some regions may still be connected through narrow *"bridging"* areas with a minimal color difference, causing large, unintended merges. To address this issue, we apply the *Louvain community detection* [BGLL08]. Within the filtered RAG, based on graph modularity, it identifies densely connected subregions separated by low connection density regions. This initially places all graph nodes in a separate region and iteratively places nodes into adjacent communities until a local modularity optimum has been reached. Nodes within the same community are merged into a single region, and the average color from the original image is assigned to the resulting abstraction.

The final abstraction reflects both the user-defined and automatically determined levels of detail. Prominent features and key boundaries are preserved, while less significant regions are merged to simplify the image meaningfully. This adaptive method enables applications in artistic rendering and robotic painting, producing abstractions that strike a balance between simplicity and detail.

4. Results

In our abstraction framework, large homogeneous regions represent objects with minimal visual or semantic impact. These include building facades, the sky, or grassy patches, which can be heavily abstracted without significantly affecting the scene's visual impression. Conversely, semantically important objects, such as people, signs, or intricate structures, are maintained with finer details, i.e. with more smaller regions in the resulting abstraction.

The computed region-based abstraction was successfully used as input for robotic painting agents implementing a polygon-filling procedure to paint the regions defined in the abstraction. This enables the creation of physical artworks with paint and canvas to study the human painting process or for artistic purposes. The infilling strategy can further profit from the computed semantic information to determine appropriate error margins during the painting process to protect sharp object boundaries and to compute brush strokes with specific directions within a region or with a different technique based on the semantic object the region is part of. Figure 2 and Figure 1e illustrate how a robotic painting system can employ our abstraction method to transform an input image into a physical painting. Our painting robot uses a DaVinci College 8 brush and water-mixable oil paints. The abstraction has been further color quantized using K-means clustering to reduce the colors to 17 for the cathedral and 12 colors for the winter house.The painting process took approximately 15 hours and 12 hours, respectively.

5. Discussion

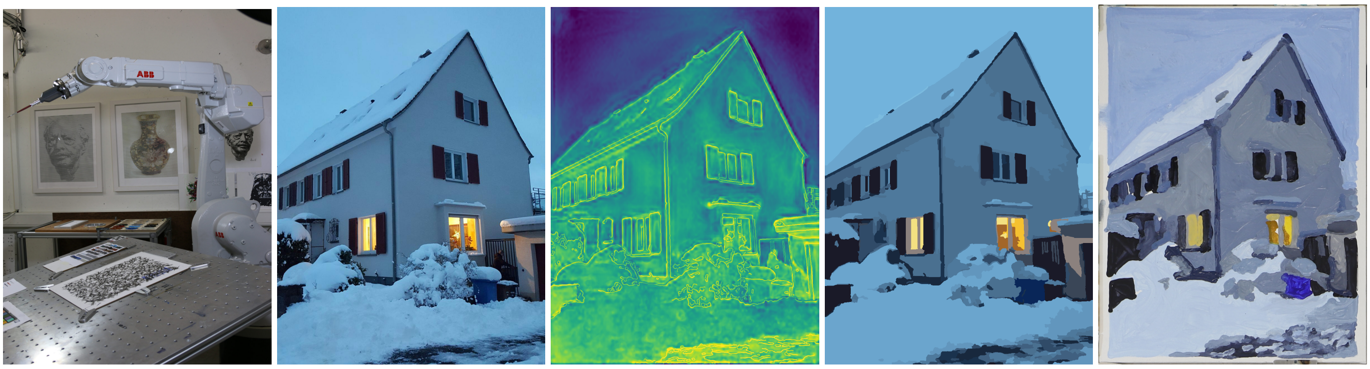
We propose an image abstraction method specifically designed to facilitate region-based robotic painting. Our method creates a semantically meaningful image abstraction with region numbers manageable by robotic painting systems, while preserving the most relevant image features without much human parameter selection for all semantic regions. We use a color-based image over-segmentation step in conjunction with panoptic segmentation to create an enhanced over-segmentation. This combination of color and panoptic segmentation resolves object boundaries without strong color differences and prevents the merging of semantically different image regions in the RAG. In this manner, we can safely avoid merging different objects, provided the segmentation model produces a good prediction on the input image.

Different methods for artistic image abstraction aim to reach diverse goals using various styles to create different aesthetic impressions. This makes quantitative comparisons between methods in this domain inherently tricky. However, a qualitative comparison is still possible. We aim to create a framework for image abstraction with object-level user guidance and automatic parameter selection that creates image regions, mimicking the human region-based painting styles. Figure 3 show results of other image abstraction methods alongside our abstractions for a qualitative comparison.

We note that the recent method for robotic painting proposed by [SGD23] requires extensive parameter tuning to balance abstraction and detail, as these parameters have to be manually adjusted for each object in the scene. In contrast, our approach automatically determines all individual parameters within a predefined abstraction range, which remains consistent across all objects in the demonstration. This automation significantly reduces the need for manual tuning, streamlining the abstraction process while maintaining high visual fidelity.

6. Conclusions

We introduced a novel framework for user-guided image abstraction designed to generate structured, semantically enriched representations well-suited for robotic painting. Our approach integrates panoptic image segmentation [KHG*19] with semantic saliency and color-based over-segmentation, enabling fine-grained control over abstraction levels. By filtering edges based on a computed threshold, determined via knee-point detection [SAIR11],



(a) **(b)** **(c)** **(d)** **(e)**

Figure 2: *Robotically painted result: b) industrial robot setup; b) original image [Ste24]; c) saliency map; d) color reduction with a palette limited to* 17 *colors to reduce the preparations necessary for robotic painting; e) robotically painted result using water-mixable oil paints.*

(a) **(b)** **(c)** **(d)**

Figure 3: *Visual Comparison: (a) original image and (b) results by Kyprianidis et al. [KD08]; (c) by Sadreazami et al. [SAM17]; and (e) by our method.*

our framework ensures that key object features and structured regions are preserved while facilitating adaptive abstraction levels per semantic object. Following edge filtering, we incorporate the Louvain community detection [BGLL08] to prevent over-merging large object regions due to thin bridging connections in the underlying Region adjacency graph structure. The resulting image abstractions are particularly beneficial for robotic painting, where controlling the level of detail, object separation, and structured abstraction is critical for guiding stroke placement and color application. Our method supports various applications, including vectorized abstraction, artistic stylization, and NPR for robotic art generation.

References

[ASS*10] ACHANTA R., SHAJI A., SMITH K., LUCCHI A., FUA P., SÜSSTRUNK S.: *Slic superpixels*. Tech. rep., 2010. 2

[BGLL08] BLONDEL V. D., GUILLAUME J.-L., LAMBIOTTE R., LEFEBVRE E.: Fast unfolding of communities in large networks. P10008. 3, 4

[FH04] FELZENSZWALB P. F., HUTTENLOCHER D. P.: Efficient graph-based image segmentation. *Int. Journ. of Computer Vision 59*, 2 (2004), 167–181. doi:10.1023/B:VISI.0000022288.19776.77. 2, 3

[FXDG17] FARAJ N., XIA G.-S., DELON J., GOUSSEAU Y.: A generic framework for the structured abstraction of images. In *Expressive - NPAR 2017* (2017), NPAR '17 Proc. of the Symp. on Non-Photorealistic Animation and Rendering. doi:10.1145/3092919.3092930. 2

[GD22] GÜLZOW J. M., DEUSSEN O.: Region-based approaches in robotic painting. doi:10.3390/arts11040077. 1

[JLC*23] JAIN J., LI J., CHIU M., HASSANI A., ORLOV N., SHI H.: OneFormer: One Transformer to Rule Universal Image Segmentation. 2

[KCWI13] KYPRIANIDIS J. E., COLLOMOSSE J., WANG T., ISENBERG T.: State of the "art": A taxonomy of artistic stylization techniques for images and video. *IEEE Trans. on Visualization and Computer Graphics 19*, 5 (2013), 866–885. doi:10.1109/TVCG.2012.160. 1

[KD08] KYPRIANIDIS J. E., DÖLLNER J.: Image abstraction by structure adaptive filtering. In *TPCG* (2008), pp. 51–58. 4

[KHG*19] KIRILLOV A., HE K., GIRSHICK R., ROTHER C., DOLLAR P.: Panoptic segmentation. In *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* (2019). 3

[LPD13] LINDEMEIER T., PIRK S., DEUSSEN O.: Image stylization with a painting machine using semantic hints. *Computers & Graphics 37*, 5 (2013), 293–301. 1, 2

[SAIR11] SATOPAA V., ALBRECHT J., IRWIN D., RAGHAVAN B.: Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st Int. Conf. on distributed computing systems workshops* (2011), IEEE, pp. 166–171. 3

[SAM17] SADREAZAMI H., ASIF A., MOHAMMADI A.: Iterative graph-based filtering for image abstraction and stylization. *IEEE Trans. on Circuits and Systems II: Express Briefs 65*, 2 (2017), 251–255. doi:10.1109/TCSII.2017.2669866. 2, 4

[SGD23] STROH M., GÜLZOW J.-M., DEUSSEN O.: Semantic image abstraction using panoptic segmentation for robotic painting. In *Proc. of Vision, Modeling, and Visualization* (2023), EG, pp. 1–8. 2, 3

[SMO23] SCHALDENBRAND P., MCCANN J., OH J.: Frida: A collaborative robot painter with a differentiable, Real2Sim2Real planning environment. In *2023 IEEE Int. Conf. on Robotics and Automation (ICRA)* (2023), pp. 11712–11718. doi:10.1109/ICRA48891.2023.10160702. 1, 2

[Ste24] STEIN M.: Winter house, 2024. 4

[TF13] TRESSET P., FOL LEYMARIE F.: Portrait drawing by paul the robot. *Computers & Graphics 37*, 5 (2013), 348–363. doi:https://doi.org/10.1016/j.cag.2013.01.012. 1

[TL12] TRESSET P. A., LEYMARIE F. F.: Sketches by paul the robot. In *Proc. of the Eighth Annual Symp. on Computational Aesthetics in Graphics, Visualization, and Imaging* (Goslar, DEU, 2012), CAe '12, Eurographics Association, p. 17–24. 1