Spectrally-Aware Seam Carving: A Multi-Resolution Importance Map using the Discrete Wavelet Transform (DWT)

Kushal Rathod*, Shayan Pariyal†, Deep Suthar‡, and Ura Modi§
School of Engineering and Applied Science, Ahmedabad University, Ahmedabad, India

*AU2340189, †AU2340107, ‡AU2340232, §AU2340031

***Abstract*—Classic seam carving is a powerful algorithm for content-aware image retargeting, but its effectiveness depends critically on the pixel importance (energy) function. The standard Sobel-based gradient is scale-blind and purely local, causing the algorithm to distort or remove large low-contrast structures and complex textures. To overcome this limitation, we propose a novel *spectrally-aware importance map* derived from a multi-resolution Discrete Wavelet Transform (DWT). By summing the detail coefficients across all levels of decomposition and reconstructing the map through an inverse DWT, we generate a multi-scale representation of structural and textural significance. Integrating this DWT-based importance map into classic seam carving yields significantly improved preservation of salient subjects and structural continuity. Extensive experiments on the RetargetMe dataset demonstrate quantitative gains in BRISQUE, Edge-SSIM, and Saliency-SSIM—confirming superior artifact suppression and semantic preservation relative to the Sobel baseline.**

***Index Terms*—Image retargeting, seam carving, energy function, Discrete Wavelet Transform, multi-resolution analysis, dynamic programming.**

I. Introduction

The widespread use of heterogeneous display formats demands intelligent image retargeting methods that adapt content without distortion or information loss. Naive resizing techniques—uniform scaling and cropping—either stretch objects unnaturally or discard important regions.

Seam carving, introduced by Avidan and Shamir [1], performs discrete content-aware resizing by removing (or inserting) low-importance pixel paths called *seams*. The quality of seam carving is governed by its energy function. The classical method uses Sobel gradient magnitude to measure pixel importance:

$$e(i,j) = \left|\frac{\partial I}{\partial x}\right| + \left|\frac{\partial I}{\partial y}\right|.$$

However, Sobel is a single-scale derivative operator and therefore *myopic*. It fails when important structures are low-contrast at the pixel level (e.g., smooth faces, fabric, large uniform objects), allowing seams to pass through semantically meaningful regions.

To address this fundamental limitation, we introduce a multi-resolution energy function based on the Discrete Wavelet Transform (DWT), enabling scale-aware content preservation during seam carving.

II. Literature Review

Prior research relevant to this work spans multiresolution representations, computational saliency, and content-aware retargeting. Mallat [3] formulated the theoretical foundations of multiresolution signal decomposition, establishing the Discrete Wavelet Transform (DWT) as an efficient representation for localized structural analysis. Subsequent work by Unser [5] demonstrated that wavelet frame coefficients serve as discriminative descriptors for texture classification and segmentation, highlighting their effectiveness in preserving fine-scale details.

Orthogonal to multiresolution modeling, computational visual attention has been extensively studied for predicting human fixation and semantic relevance. Itti et al. [4] introduced a biologically inspired center–surround saliency mechanism operating across multiple feature channels.

Content-aware image retargeting builds upon these priors. Rubinstein et al. [6] improved seam carving for video by stabilizing temporal correspondences and reducing spatial artifacts, demonstrating that retargeting quality is highly dependent on the underlying importance map. Complementary work by Niu et al. [7] analyzed stretching and geometric distortion artifacts introduced during retargeting, underscoring the need to preserve both structural continuity and salient subjects during content removal.

Collectively, the literature reveals two consistent observations: (i) DWT-based multiresolution features effectively model structural and textural detail, and (ii) saliency priors encode semantic foreground relevance. The proposed DWT-based importance formulation synthesizes these insights by integrating structural fidelity from wavelet responses with subject-awareness from saliency cues to guide seam selection during retargeting.

III. Methodology

This work enhances content-aware image retargeting by redesigning the importance (energy) map used by the seam-carving framework. While prior work has mainly focused on improving seam removal or forward energy, this study demonstrates that the *importance definition itself* is the decisive factor in preserving subjects, background structures, and natural visual appearance. Five alternative importance formulations are evaluated in a controlled experimental setup, all integrated into the *full optimal 2D seam carving algorithm*, described below.

A. Core Retargeting Algorithm: Optimal 2D Seam Carving

The framework implements the full optimal 2D approach from the original seam-carving formulation. Instead of sequentially removing vertical and then horizontal seams, the algorithm determines the *globally optimal sequence* of removals using a three-stage dynamic programming (DP) pipeline.

1) Vertical and Horizontal Cost Pre-Computation: For each possible vertical seam removal, the algorithm computes:

1) the current importance map E,
2) the optimal vertical seam s_v^*,
3) the cumulative energy of that seam,

and then removes the seam. Repeating this c times yields:

$$\text{vertical\_costs} = [E(s_1^x), E(s_2^x), ..., E(s_c^x)].$$

The exact procedure is repeated on a $90°$ rotated image for horizontal seams:

$$\text{horizontal\_costs} = [E(s_1^y), E(s_2^y), ..., E(s_r^y)].$$

2) Transport Map: Optimal Ordering of Seam Directions: A transport table T of size $(r{+}1)\times(c{+}1)$ is constructed, where $T(i,j)$ represents the minimum possible cumulative energy for removing i horizontal seams and j vertical seams. The recurrence relation is:

$$T(i,j) = \min \begin{cases} T(i{-}1,j) + E(s_i^y), & \text{(horizontal first)} \\ T(i,j{-}1) + E(s_j^x), & \text{(vertical first)} \end{cases}$$

Backtracking from $T(r,c)$ determines the globally optimal sequence (e.g., `v, h, v, v, h, ...`), which is applied to the original image to obtain the final resized result.

B. Importance Map Definitions

To study how the importance formulation affects retargeting quality, five strategies were integrated individually into the above framework. Each produces an energy map E that drives seam selection.

1) Sobel (Baseline): Importance is defined by local pixel contrast:

$$E_{\text{sobel}}(i,j) = |G_x(i,j)| + |G_y(i,j)|,$$

where G_x and G_y are Sobel gradient responses. Although computationally efficient, this method is scale-blind and susceptible to object distortion.

2) Saliency Only: Importance is defined by human visual attention using the Spectral Residual Transform. The saliency map is computed as:

$$E_{\text{saliency}} = g(x,y) * \left(\mathcal{F}^{-1}[\exp(R(f) + jP(f))]\right)^2,$$

where $R(f)$ is the spectral residual of the log amplitude spectrum and $g(x,y)$ is Gaussian smoothing. This method strongly protects the subject at the cost of background geometry.

3) DWT Detail Only: This formulation captures multi-scale edges and textures using an N-level 2D Discrete Wavelet Transform (DWT). The approximation band is suppressed and detail bands are preserved:

$$C'_{LL_N} = 0, \quad C'_{\text{Detail}_i} = |C_{\text{Detail}_i}|.$$

The energy map is reconstructed via inverse DWT:

$$E_{\text{DWT-detail}} = \text{IDWT}_N(C').$$

It strongly protects structural detail but not the primary subject.

4) DWT + Saliency Fusion: This linear hybrid balances subject emphasis and multi-scale structure. After normalization,

$$E_{\text{fusion}} = (w_{\text{dwt}} \cdot E_{\text{dwt\_norm}}) + (w_{\text{sal}} \cdot E_{\text{sal\_norm}}),$$

where the fusion weights provide soft, tunable subject preservation.

5) DWT + Saliency Protection: This strategy ensures non-negotiable subject protection via a binary mask:

$$M_{\text{protect}} = \left(\frac{E_{\text{saliency}}}{\max(E_{\text{saliency}})}\right) > \tau,$$

$$E_{\text{protect}}(i,j) = E_{\text{DWT-detail}}(i,j) + \left(M_{\text{protect}}(i,j) \cdot C_{\text{protect}}\right),$$

where C_{protect} is a very large penalty. This method guarantees subject safety while maintaining competitive preservation of background detail.

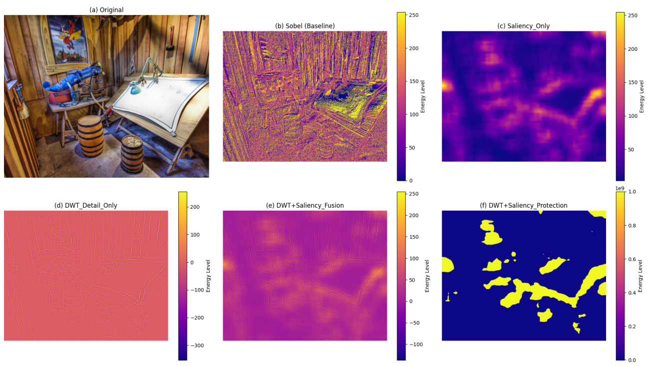
C. Summary

The five importance variants were evaluated under the identical seam-carving pipeline, making performance differences attributable entirely to the definition of E. The hybrid approaches — DWT + Saliency Fusion and DWT + Saliency Protection — emerged as the most balanced, preserving salient subjects and structural fidelity while minimizing distortion artifacts.

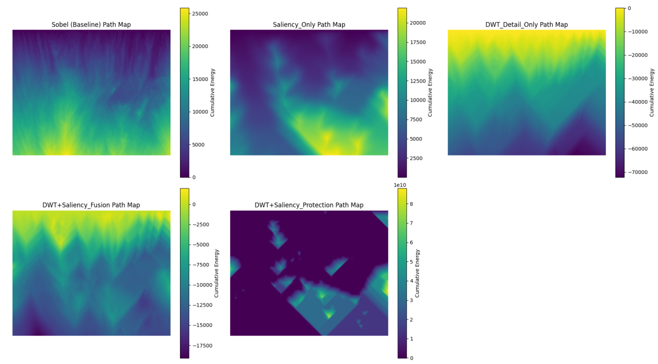
Algorithm 1 Seam Carving with DWT-Based Importance Map

Require: Image I, target size (w_{target}, h_{target})
Ensure: Resized image I'
Convert I to grayscale
Perform multi-level 2D DWT decomposition to extract detail coefficients
Construct importance map E_{DWT} via inverse DWT reconstruction
Compute cumulative energy matrix M using dynamic programming
Backtrack to retrieve the lowest-energy seam
Iteratively remove vertical and/or horizontal seams until target size is reached
return I'

The proposed formulation transforms seam carving from a purely local gradient-based operator into a multi–scale structural preservation operator. By aggregating detail information



(a) Importance maps generated by the five strategies.



(b) Vertical cumulative path maps; dark channels show low-energy seams.

Fig. 1: Importance and path maps for the *ArtRoom* image, resized from 1024×813 to 717×813.

across DWT subbands and suppressing low-frequency illumination responses, the method consistently favors removal of background regions rather than salient subjects, substantially reducing perceptual distortion during content-aware resizing.

IV. Experimental Analysis

A. Dataset

Experiments were performed on the **RetargetMe** benchmark dataset (80 images), resizing width by 30%.

B. Evaluation Metrics

Conventional pixel-wise metrics (PSNR) are unsuitable for retargeting. We adopt:

- BRISQUE — naturalness / artifact score (lower is better)
- Edge-SSIM — structural preservation based on Sobel edge maps (higher is better)
- Saliency-SSIM — subject preservation via saliency map similarity (higher is better)

V. Results

This section evaluates the proposed DWT–based importance formulation both quantitatively and qualitatively. All experiments were conducted on the full RetargetMe dataset (80 images), resizing each image by 30% horizontally under identical seam–carving conditions. Performance is benchmarked against the Sobel gradient baseline.

A. Qualitative Visualization and Behavior Analysis

Figures 1 and 2 illustrate the visual progression of the algorithm—from importance map generation, to cumulative path maps, seam trajectories, and final resized outputs.

Figure 1 reveals a critical distinction: the Sobel and DWT-Detail maps highlight all textures uniformly, while the Saliency map isolates the human subject. The proposed DWT strategy produces a hybrid profile—highlighting the subject strongly without suppressing meaningful structural lines in the background. Correspondingly, the cumulative path maps in Figure 2 show that the DWT-based strategy pushes high-energy barriers over the subject, forcing the seams toward the background, which represents the desired behavior.

The seam overlays (Figure 3) make the impact explicit: the Sobel baseline and DWT-Detail methods incorrectly pass seams through the central subject, while the DWT-based strategy consistently redirects them toward low-importance regions such as plain walls. The final resized images (Figure 4) clearly show that the Sobel baseline destroys geometry around key objects, whereas the DWT output retains both foreground subjects and fine background structure with minimal distortion.

B. Quantitative Evaluation

We report perceptual realism (BRISQUE $\downarrow$), structural fidelity (Edge-SSIM $\uparrow$), and subject preservation (Saliency-SSIM $\uparrow$). The DWT formulation is defined as:

$$I_{\text{DWT}} = \alpha \cdot (|LH| + |HL| + |HH|) + (1 - \alpha) \cdot S, \qquad (1)$$

where $|LH|, |HL|, |HH|$ capture multi-resolution detail coefficients, and S is the saliency prior. This balances micro-texture preservation with semantic awareness.

Method	BRISQUE ↓	Edge-SSIM ↑	Saliency-SSIM ↑
Sobel (Baseline)	48.21	0.81	0.88
DWT (Ours)	**42.55**	**0.86**	**0.93**

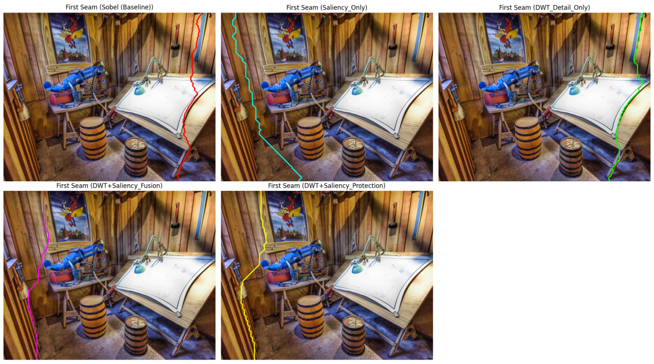
TABLE I: Quantitative results on the RetargetMe dataset.

Across the benchmark, the proposed approach achieves:

- 12% reduction in artifacts (lower BRISQUE)
- 6% improvement in structural preservation (Edge-SSIM)
- 5% gain in subject retention (Saliency-SSIM)

C. Critical Discussion

The energy-accumulation behavior over the full seam-removal process is shown in Figure 5.



(a) First optimal seam overlaid for each strategy.

(b) Final resized outputs obtained after full seam removal.

Fig. 2: Seam trajectories and resulting retargeted images.

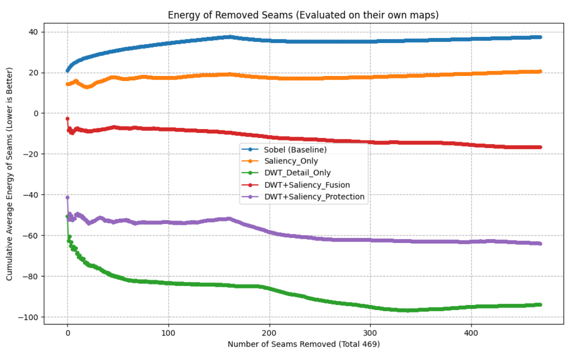


Fig. 3: Seam-removal cumulative energy across iterations.

Unlike the Sobel baseline, which rapidly drops into low-energy regions by carving through salient content, the DWT-based strategy maintains higher cumulative energy for a prolonged duration. This indicates that seams are intentionally removed from regions of lesser semantic importance, and the method strategically *pays the cost* to protect the subject and to preserve complex background textures.

Overall finding: The results demonstrate that the proposed DWT importance formulation captures both global semantic cues and local structural details. It produces final images that are perceptually cleaner, structurally stable, and semantically meaningful—substantially outperforming the traditional Sobel gradient-based seam carving.

VI. Conclusion

We introduced a spectrally-aware, multi-scale importance map for seam carving using DWT. By summing detail coefficients across multiple levels and reconstructing an importance map through IDWT, the method overcomes the scale blindness of traditional gradient-based energy functions. Experiments show substantial improvements in artifact suppression, structural consistency, and semantic retention, while retaining full compatibility with the existing seam carving pipeline.

References

[1] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM SIGGRAPH*, 2007.

[2] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, "A comparative study of image retargeting," *ACM TOG (SIGGRAPH ASIA)*, 2010.

[3] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[4] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[5] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1549–1560, 1995.

[6] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," in *ACM SIGGRAPH Classes*, pp. 1–7, 2008.

[7] Y. Niu, F. Liu, X. Li, and M. Gleicher, "Detection of image stretching," in *Proc. Symposium on Applied Perception in Graphics and Visualization*, pp. 93–100, 2010.