Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 22, No. 1, 2026

A Data-Driven Closed-Loop Parameter Adaptive Framework for Automated Visual Standardization Systems

Guozhi Hu[1, 2] * (0009-0006-7282-2072) and Robiatul A'dawiah Jamaluddin[1] (0009-0006-2617-0069)

[1] Kuala Lumpur University of Science and Technology (KLUST), Kuala Lumpur, Malaysia
[2] Beijing Kingsha Technology Co., Ltd., Beijing, China

*** Corresponding author:** Guozhi Hu (242924843@s.klust.edu.my)

Abstract: Automated visual standardization — the conversion of arbitrary user-provided portraits into images that satisfy strict format specifications — is a key component of high-reliability identity-management systems. Conventional pipelines operate as feedforward open-loop systems: they map facial landmarks to a fixed scaling ratio and synthesize the output in a single pass, implicitly assuming a linear relationship between facial scale and body anatomy. Under high-dimensional stochastic geometric variation (narrow or wide shoulders, long hair, atypical clothing), this assumption breaks down and produces shoulder truncation or bottom-edge gaps that violate format compliance. This paper reformulates visual standardization as a feedback-controlled parameter-optimization process. Our framework cascades perception-aware color normalization, BiRefNet-based foreground segmentation with distance-transform edge purification, and a closed-loop geometric adaptation stage in which an edge-integrity sensor measures posterior canvas coverage and a controller iteratively refines the affine scale parameter until convergence. On 120 real-world samples spanning 12 challenging categories, the closed-loop framework raises the bottom-edge touch rate from 99.17% to 100% and bottom-row foreground coverage from 98.73% to 99.99%, while reducing the boundary contamination ratio from 67.5% to 53.8%. Ablation studies confirm that the geometric feedback loop is the sole factor eliminating bottom-edge gaps and that edge purification independently accounts for a 15.9-percentage-point reduction in boundary contamination. The results demonstrate that posterior feedback converts residual geometric uncertainty into a convergent control problem, enabling deterministic compliance under strict business constraints.

Keywords: Affine transformation; Closed-loop control; Edge integrity; Parameter adaptation; Portrait segmentation; Visual standardization.

1. Introduction

In government services, examination registration, recruitment, and online identity verification, automated systems must convert unconstrained user-uploaded portraits into outputs that satisfy rigid specifications: fixed pixel dimensions, uniform background color, prescribed face-width ratio, level head pose, and — critically — a foreground that reaches and fully covers the bottom edge of the canvas. The input to such systems is doubly uncertain: photometrically, because of random illumination, color casts, and exposure defects; and geometrically, because of the stochastic variation of human posture, body shape, hairstyle, and clothing.

Conventional standardization strategies address only the first-order geometry: a face detector provides landmarks, a preset face-width ratio determines a fixed scale factor, and the output is composed in a single feedforward pass. Such open-loop designs assume that body anatomy scales linearly with facial dimensions. In practice this assumption fails systematically: narrow-shouldered subjects leave uncovered corners on the bottom row, and long hair that extends below the crop boundary can leave the entire bottom rows empty. Because the open-loop system has no mechanism to observe its own output, these violations pass through to the final product and cause compliance failures that require manual rework.

This paper recasts visual standardization as a closed-loop control problem, in which the synthesized canvas itself is measured and the geometric parameters are iteratively corrected until the output satisfies the format constraints.

Specifically, we first stabilize the photometric properties of the input with a perception-aware color-normalization stage built on facial statistical profiling and spatially variant gain control, so that the segmentation stage receives consistent inputs across lighting conditions. We then define an edge-integrity feedback sensor on the alpha channel of the target canvas: a posterior evaluation function that measures bottom-edge touch and bottom-row foreground coverage, converting geometric compliance into a measurable deviation signal. Around this sensor we build a closed-loop parameter adaptation — an iterative search over the affine scale, driven by the feedback signal and terminated by explicit convergence criteria — that guarantees geometric compliance, while a single-pass affine synthesis avoids cumulative resampling degradation.

2. Related Work

2.1. Automated portrait standardization

Existing automated ID-photo pipelines are predominantly feedforward: landmark-driven cropping followed by background replacement. Photometric preprocessing in such systems draws on classical enhancement theory — Retinex [11], multiscale retinex with color restoration [12], illumination-map estimation [13], and more recently zero-reference deep curve estimation [14]. In our framework, photometric normalization follows the perception-aware statistical approach described in a companion study [15], which balances brightness recovery, highlight suppression, and skin-tone naturalness through interpretable constraints; the present paper focuses on the geometric and boundary-integrity dimensions of the standardization problem.

2.2. Portrait segmentation and matting

High-quality foreground extraction is the prerequisite of background replacement. Salient-object and dichotomous segmentation networks such as U2-Net [2] and BiRefNet [1] provide high-resolution alpha priors, and trimap-free matting networks such as MODNet [3] target real-time portrait applications. The composition of a segmented foreground onto a new background follows the classical alpha-compositing algebra of Porter and Duff [4], while closed-form matting [5] established the analytic treatment of semi-transparent boundaries. A persistent practical issue is boundary contamination: semi-transparent hair and clothing edges retain colors of the original background. Our purification mechanism addresses this with a distance-transform-based color-borrowing scheme [6, 7] that re-synthesizes fringe pixels from the nearest solid foreground.

2.3. Feedback control in vision pipelines

Closed-loop regulation is a foundational concept in control engineering [8]: a sensor measures the system output, the deviation from a reference is fed back to a controller, and the actuating parameter is adjusted until the deviation converges. Vision pipelines, by contrast, are usually built as open-loop cascades, with each stage trusting the output of its predecessor. Our framework transfers the feedback principle to geometric standardization by treating the affine scale as the actuated variable and posterior canvas measurements as the sensor signal. Single-pass affine synthesis, motivated by the resampling analysis of digital image warping [9], prevents the aliasing accumulation that would arise from iteratively re-warping intermediate results; face localization uses a standard detector [10].

3. Problem Formulation

Visual standardization maps an input image I to an output canvas O of fixed dimensions W × H with a specified background color, face-width ratio ρ, and top-margin ratio m. The mapping is driven by an affine transformation:

$$O = T(I, P), \quad P = \{s, \theta, tx, ty\}, \tag{1}$$

where s is the scale factor, θ the rotation angle estimated from ocular landmarks, and (tx, ty) the translation vector. The output must simultaneously satisfy three families of constraints: photometric constraints (natural facial brightness and skin tone, uncontaminated background), boundary constraints (continuous alpha edges that blend with the target background), and geometric constraints (face-width ratio, top margin, and complete bottom-edge closure). The task is to find the parameter set P for which all constraints hold. Open-loop systems fix s from facial geometry alone; the framework proposed below makes s the subject of a feedback-controlled search.

4. The Proposed Framework

Figure 1 shows the architecture. The forward path cascades three stages — perceptual color normalization, foreground segmentation with edge purification, and single-pass affine synthesis. The feedback path projects the alpha channel of the synthesized canvas into an edge-integrity sensor whose deviation signal drives a scale controller; the loop iterates until the geometric constraints are met.

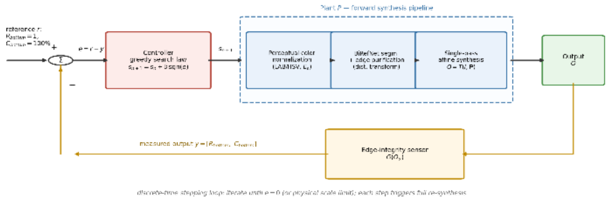

Figure 1. Closed-loop architecture of the proposed framework in standard control-system notation: reference r (full bottom-edge closure) and measured output y = [Rbottom, Cbottom] form the error e = r − y at the summing junction; the controller applies a greedy stepping law to the actuated scale variable s, which drives the forward synthesis plant P; the edge-integrity sensor G(Oα) closes the loop.

4.1. Perceptual color normalization

The photometric front-end stabilizes the input signal before segmentation. Facial statistical profiling extracts L-channel percentiles (p10–p95) in the LAB space to form a robust exposure anchor; color-cast correction operates on truncated a/b means through a dead-zone-and-damping rule that preserves natural skin warmth while removing environmental casts; and local compensation is applied through a face-centric Gaussian soft mask rather than a global gain, preventing background saturation in backlit or white-background scenes. The complete formulation and its dedicated evaluation are given in the companion study [15]; within this paper the stage serves as spatially variant gain control that normalizes the input distribution for the downstream stages.

4.2. Foreground segmentation and edge purification

BiRefNet [1] generates a high-resolution alpha matte, which is refined by an S-curve transfer that steepens the transition of semi-transparent regions. Residual boundary contamination — original-background color retained in hair fringes and collar edges — is then removed by a purification mechanism: (i) a solid-foreground mask is extracted by thresholding the alpha matte; (ii) a distance transform [6, 7] identifies, for every fringe pixel, its nearest solid-foreground pixel; (iii) the fringe pixels color is re-synthesized by borrowing the solid pixels chromatic components and blending them with the target background according to the enhanced alpha value:

$$Cedge = \alpha \cdot Csolid + (1 - \alpha) \cdot Cbg. \tag{2}$$

This guarantees that semi-transparent boundaries converge to the target background color rather than to the colors of the original scene, following the compositing algebra of [4].

4.3. Closed-loop adaptive geometric standardization

Feedback sensor. After each synthesis, the edge-integrity function $G(O_\alpha)$ inspects the alpha channel of the output canvas and reports two posterior measurements: the bottom-edge touch indicator R_{bottom} (whether foreground reaches the last canvas row) and the bottom-row coverage C_{bottom} (the fraction of bottom-row pixels occupied by foreground). Coverage below 100% indicates a geometric gap.

Controller. The scale factor s is the actuated variable, initialized from the preset face-width ratio (s_0). Denoting the reference as r (full bottom-edge closure) and the sensor measurement as y, the error signal $e = r - y$ drives a greedy search control law — a discrete-time stepping controller acting on the scale variable:

$$s(n+1) = s(n) + \delta \cdot \mathrm{sgn}(e), \tag{3}$$

where δ is the search step. When a gap is detected below the shoulders, the law yields an enlargement step; for narrow-body subjects whose torso must enter the frame, the body-entry branch yields a reduction step. Each update triggers a complete re-synthesis, so the sensor always measures the true output rather than a prediction. **Convergence.** The iteration terminates when $R_{\text{bottom}} = 1$ and $C_{\text{bottom}} = 100\%$, or when a physical scale limit is reached, at which point the best parameter set found is used. Because the actuated variable is one-dimensional and the constraint set is monotone in s within the search interval, the loop converges in a small, bounded number of iterations and cannot oscillate.

Single-pass synthesis. The final rotation, scaling, and translation are combined into one affine matrix and applied in a single resampling operation. This avoids the cumulative aliasing and detail loss of sequential rotate–resize–paste implementations [9], and ensures that iterative search does not degrade image quality: intermediate iterations only measure the canvas, and the output is always synthesized directly from the source image.

5. Experimental Design and Analysis

5.1. Dataset and setup

Evaluation uses 120 real-world samples across 12 categories in six scenario groups: lighting (low light, overexposure), color deviation (yellow indoor light, reddish skin tone), background complexity (complex, light), clothing color (dark, light), body geometry (wide shoulders, narrow shoulders), and hair boundary (long hair, wispy hair). Four system configurations are compared, all sharing identical face detection and BiRefNet segmentation:

B0 — feedforward baseline: no color processing, fixed face-width scaling;

B1 — simple global brightness/contrast enhancement, fixed scaling;

B2 — full color normalization, fixed scaling (open-loop);

Proposed — full color normalization with edge purification and the closed-loop geometric adaptation.

Three families of metrics are reported. Photometric quality is summarized by the central L-mean and the highlight ratio (share of pixels with L > 245). Geometric compliance is measured by the bottom-edge touch rate (share of samples whose foreground reaches the last canvas row) and the mean bottom-row foreground coverage. Boundary purity is measured by the edge contamination ratio: the share of foreground-boundary pixels whose color deviates from the target background beyond a fixed threshold.

5.2. Overall results

Table 1 aggregates the results over all 120 samples. All fixed-scale configurations (B0, B1, B2) share identical geometric failures — a 99.17% touch rate and 98.7% bottom-row coverage — regardless of their photometric quality, confirming that geometric compliance is orthogonal to color processing and cannot be recovered by enhancement alone. The closed-loop framework achieves 100% touch rate and 99.99% mean coverage while simultaneously reducing edge contamination from 67.5% to 53.8%.

Table 1. Overall performance over 120 samples.

Method	L-mean	HL (%)	Touch (%)	Cover (%)	Edge (%)
B0	144.64	0.42	99.17	98.73	67.54
B1	160.10	3.00	99.17	98.72	67.93
B2	165.89	0.70	99.17	98.72	67.55
Ours	165.55	0.72	100.00	99.99	53.78

5.3. Failure-case analysis

Figure 2 examines the two characteristic failure modes of open-loop scaling on the bottom band of the output canvas. In the long-hair case (EXP110), fixed scaling leaves a 24-pixel horizontal gap between the foreground and the canvas bottom — bottom-row coverage drops to 0% and the sample fails compliance outright; the closed-loop search enlarges the scale until the gap closes completely. In the narrow-shoulder case (EXP092), the shoulders do not span the canvas width at the bottom row, leaving a corner notch with 95.4% coverage; the feedback loop resolves it to 100%. Table 2 summarizes both cases.

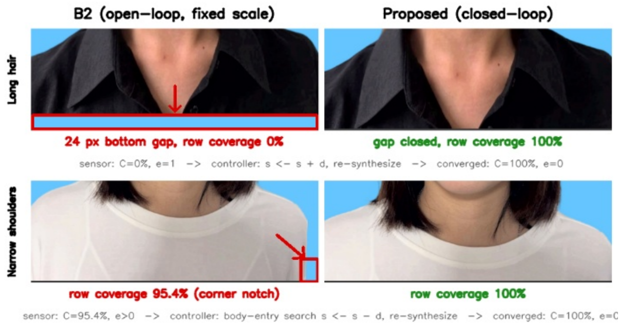

Figure 2. Bottom-band views (no faces shown) of the two open-loop failure modes and their closed-loop corrections. Top: long-hair sample with a 24 px bottom gap. Bottom: narrow-shoulder sample with a bottom-row corner notch. Red marks indicate the deviation regions detected by the edge-integrity sensor; the grey line under each case gives the sensor–controller sequence that closes the loop.

Table 2. Case analysis of geometric failures.

Sample	Metric	B2	Proposed
EXP110 (hair)	Gap (px)	24	0
EXP110 (hair)	Cover (%)	0.0	100.0
EXP092 (nrw.)	Cover (%)	95.4	100.0

5.4. Ablation study

Table 3 isolates the two mechanisms. Removing edge purification (A3) leaves geometric compliance intact but raises edge contamination from 53.8% to 69.7% — worse than the open-loop baseline, since adaptive enlargement magnifies contaminated fringes. Freezing the geometric search (A4) preserves the purification benefit (51.6% contamination) but reverts geometric compliance exactly to the open-loop failure level (99.17% touch, 98.72% coverage), demonstrating that the feedback loop is the sole factor eliminating bottom-edge gaps. Figure 3 visualizes the purification effect on a long-hair boundary and the corresponding contamination statistics.

Table 3. Ablation of the two core mechanisms.

Variant	Touch (%)	Cover (%)	Edge (%)
A3 (no purif.)	100.00	99.99	69.71
A4 (fixed scale)	99.17	98.72	51.56
Ours	100.00	99.99	53.78

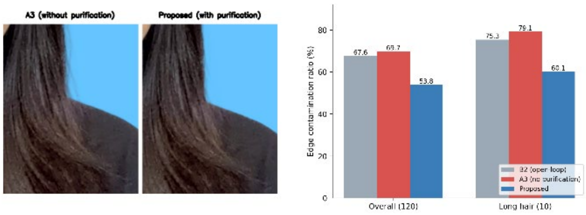

Figure 3. Edge purification. Left: hair-boundary crops (no faces shown) without and with purification — the contaminated fringe converges to the target background color. Right: edge contamination ratio overall and in the long-hair category.

5.5. Stability under nominal inputs

An auxiliary set of 24 volunteer portraits captured under controlled geometric conditions (four lighting scenarios, standard posture) was processed by all configurations. Every method, including the fixed-scale baselines, achieved 100% geometric compliance on this set, and the closed-loop framework produced identical geometry to the open-loop configuration: when the deviation signal is zero at the first iteration, the controller performs no correction. This confirms that the feedback loop activates only when a violation is detected and introduces no perturbation — and in particular no oscillation — on nominal inputs.

6. Discussion

The experiments show that the closed-loop mechanism absorbs precisely the geometric uncertainties — narrow shoulders, long hair extending past the crop boundary — that cause compliance failures in open-loop systems. A single deep model inferring the output directly offers no guarantee that a hard business constraint such as full bottom-edge closure is met on any specific input; the feedback design provides this guarantee constructively. Because the sensor measures the actual synthesized canvas, even noisy upstream stages (imperfect landmarks, slightly conservative alpha mattes) are compensated by the iterative correction, and the terminal state satisfies the constraint whenever it is physically attainable within the scale limits. The cost is bounded re-synthesis: on our data the search terminates within a few iterations, and single-pass synthesis keeps the final quality independent of how many iterations were needed.

The approach is not without boundaries. Samples whose torso is severely truncated in the source photograph cannot be repaired by scaling alone — the physical scale limit then terminates the search with a best-effort output — and the current sensor observes only alpha-channel geometry, so semantic anomalies such as collar asymmetry remain invisible to it. Both point in the same direction: introducing a human-parsing model as an additional semantic sensor would extend the feedback loop from geometric closure to semantic-level constraints on neckline and shoulder contours.

7. Conclusion

This paper reformulated automated visual standardization from a feedforward cropping task into a feedback-controlled parameter-optimization process. The framework combines perception-aware color normalization, distance-transform edge purification, and a closed-loop geometric adaptation whose sensor–controller loop iterates the affine scale until posterior edge-integrity constraints are satisfied. On 120 challenging real-world samples the framework achieved 100% bottom-edge touch and 99.99% bottom-row coverage — eliminating the compliance failures inherent to fixed-scale designs — while reducing boundary contamination from 67.5% to 53.8%. Ablations attribute each gain to its dedicated mechanism. The results support a broader design principle for high-reliability vision systems: hard output constraints are best enforced not by ever-stronger forward predictors, but by measuring the output and closing the loop.

Ethics Statement

The 120-image evaluation dataset originates from a commercial portrait-standardization service under user agreements permitting research use; it is employed for aggregate statistical evaluation only. Figure crops from this dataset (Figures 2 and 3) are restricted to shoulder, clothing, and hair-boundary regions and contain no identifiable faces. The auxiliary 24-image set depicts volunteers who provided written informed consent.

References

[1] Zheng, P., Xu, J., Liu, Z., Jiang, Y., Chen, Y., & Fan, H. (2024). Bilateral reference for high-resolution dichotomous image segmentation. CAAI Artificial Intelligence Research, 3, Article 9150038. https://doi.org/10.26599/AIR.2024.9150038

[2] Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., & Jagersand, M. (2020). U2-Net: Going deeper with nested U-structure for salient object detection. Pattern Recognition, 106, Article 107404. https://doi.org/10.1016/j.patcog.2020.107404

[3] Ke, Z., Sun, J., Li, K., Yan, Q., & Lau, R. W. H. (2022). MODNet: Real-time trimap-free portrait matting via objective decomposition. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 1, pp. 1140-1147). AAAI Press. https://doi.org/10.1609/aaai.v36i1.19998

[4] Porter, T., & Duff, T. (1984). Compositing digital images. ACM SIGGRAPH Computer Graphics, 18(3), 253-259. https://doi.org/10.1145/964965.808606

[5] Levin, A., Lischinski, D., & Weiss, Y. (2008). A closed-form solution to natural image matting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2), 228-242. https://doi.org/10.1109/TPAMI.2007.1177

[6] Rosenfeld, A., & Pfaltz, J. L. (1966). Sequential operations in digital picture processing. Journal of the ACM, 13(4), 471-494. https://doi.org/10.1145/321356.321357

[7] Felzenszwalb, P. F., & Huttenlocher, D. P. (2012). Distance transforms of sampled functions. Theory of Computing, 8(1), 415-428. https://doi.org/10.4086/toc.2012.v008a019

[8] Åström, K. J., & Murray, R. M. (2008). Feedback systems: An introduction for scientists and engineers. Princeton University Press.

[9] Wolberg, G. (1990). Digital image warping. IEEE Computer Society Press.

[10] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Vol. 1, pp. 511-518). IEEE. https://doi.org/10.1109/CVPR.2001.990517

[11] Land, E. H. (1977). The Retinex theory of color vision. Scientific American, 237(6), 108-128. https://doi.org/10.1038/scientificamerican1277-108

[12] Jobson, D. J., Rahman, Z., & Woodell, G. A. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. IEEE Transactions on Image Processing, 6(7), 965-976. https://doi.org/10.1109/83.597272

[13] Guo, X., Li, Y., & Ling, H. (2017). LIME: Low-light image enhancement via illumination map estimation. IEEE Transactions on Image Processing, 26(2), 982-993. https://doi.org/10.1109/TIP.2016.2639450

[14] Guo, C., Li, C., Guo, J., Loy, C. C., Hou, J., Kwong, S., & Cong, R. (2020). Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1780-1789). IEEE. https://doi.org/10.1109/CVPR42600.2020.00185

[15] Hu, G., & Jamaluddin, R. A. (2026). An adaptive portrait image enhancement algorithm based on facial and scene statistics in complex lighting. Manuscript submitted for publication.