Art from Algorithms: Saliency-Guided Digital Projections

Sara Greenberg University of Waterloo, ON, Canada
Alexander Wong University of Waterloo, ON, Canada

Abstract

Saliency algorithms provide a measure of which sections in an image are important or informative. This can be valuable to digital artists who wish to produce works that build layers of abstraction or augmentation onto an input image. Using a framework based on the Texture-Illumination Guided Global Energy Response (TIGGER) to access two types of saliency measurements, a multidisciplinary projection project, Nebula, has been produced, in which starscapes are created in near real time from depth images.

1 Introduction

Abstracting or augmenting a raw image is a common practice in various artistic disciplines, such as photo manipulation and projection mapping. Often, the image features of interest are identified manually. Saliency is a measure of the level of importance of image features, and therefore we propose a framework for automatic artistic image abstraction built on a saliency algorithm.

The proposed framework is based off of the Texture-Illumination Guided Global Energy Response (TIGGER) [1] for illumination robust global saliency. The components of interest of the TIGGER are the multi-scale evaluation, which provides a measure of salience for a range of feature sizes, and the final energy response provides a single measure for the entire image. These components provide a rich interpretation of the "important" parts of an image, which can then be mapped to new, creative imagery. The TIGGER provides a mapping for an RGB image, but it is possible to bypass the texture-illumination decoupling and pass in any array of values.

The flow of such a system is described by example in Fig. 1, where a performance is augmented by capturing the scene using a depth camera, processing the image to acquire the energy response and saliency metrics, applying a form of graphical mapping, and finally, re-projecting the result back onto the scene.

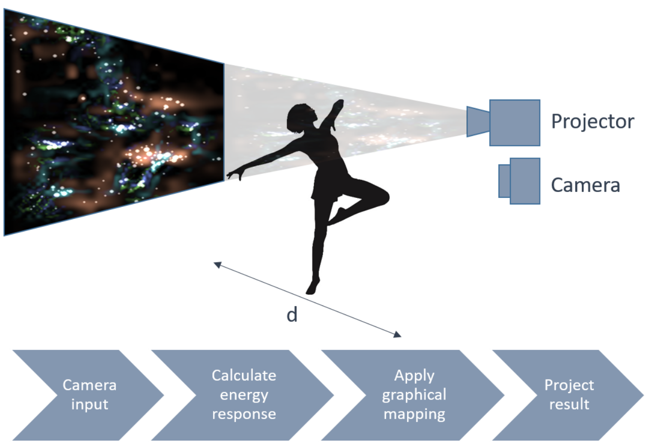

Fig. 1: Example use of multi-scale saliency and the aggregate energy response for live projection-augmented performance.

2 Approach

The proposed framework consists of two mapping methods: multi-scale saliency and the aggregate energy response. For both methods the saliency is calculated as an energy response as described by Greenberg et. al [1].

2.1 Multi-Scale Saliency

At each scale factor λ, the saliency metric is computed using the Hessian matrix, Φ, obtained for each pixel, $\bar{q}$, of an image, I, as

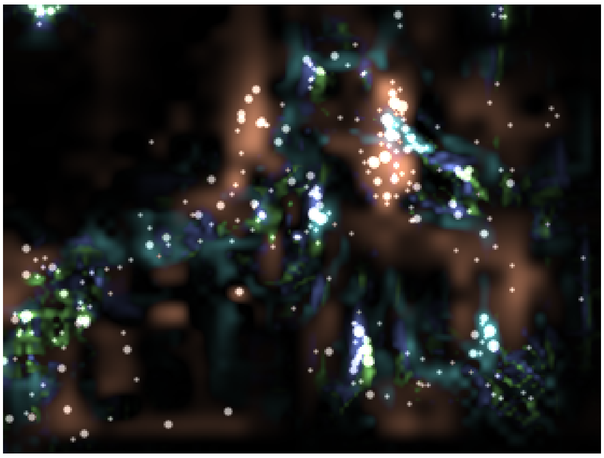

Fig. 2: Calculating the energy response of a depth image of a person standing in front of the camera at 3 scales and randomly assigning a colour to the result of each scale response.

$$\Phi(\bar{q},\lambda) = \begin{bmatrix} \left(\Delta_{x,\lambda} I(\bar{q})\right)^2 & \Delta_{x,\lambda} I(\bar{q}) \Delta_{y,\lambda} I(\bar{q}) \\ \Delta_{y,\lambda} I(\bar{q}) \Delta_{x,\lambda} I(\bar{q}) & \left(\Delta_{y,\lambda} I(\bar{q})\right)^2 \end{bmatrix} \tag{1}$$

where Δ_x and Δ_y are gradients in the x and y directions, respectively. The saliency s at scale λ of pixel $\bar{q}$ is:

$$s(\bar{q},\lambda) = \frac{det(\Phi(\bar{q},\lambda))}{trace(\Phi(\bar{q},\lambda))} \tag{2}$$

Abstraction or mapping can then be applied to s at each scale.

2.2 Aggregate Energy Response

To capture salient characteristics of the image as a whole, the aggregate response Θ is calculated as:

$$\Theta(\bar{q}) = \sum_{\lambda\in\Lambda} \frac{det(\Phi_\tau(\bar{q},\lambda))}{trace(\Phi_\tau(\bar{q},\lambda))}. \tag{3}$$

To randomly sample only points in the image with an energy response above some threshold, we implement acceptance-rejection sampling as described in algorithm 1.

Algorithm 1: Acceptance-rejection sampling for a pixel $\bar{q}$.

count = 0 ;
while *count < desired samples* **do**
  if $\Theta(\bar{q})$ > *threshold* **and** $\Theta(\bar{q})$ > *rand* **then**
    accept $\bar{q}$;
    count++ ;

3 Nebula

The Nebula project uses the proposed framework to produce space-inspired imagery that maintains characteristics of an input depth image. It uses both the individual scaled responses and the final aggregate response to create starscapes.

Each of the individual scaled responses is assigned a random colour and rescaled to the original image size. This results in different levels of saliency represented with a cloud of colour. To create the appearance of a cluster of stars, bright circles are assigned at random points in the image that have an aggregate response value above a threshold. The radius of these circles is proportional to the magnitude of the response value at the corresponding depth image pixel. The result (Fig. 2) is an image that maintains some of the shape and features of the input depth image, but is abstracted to resemble a nebula, galaxy, or other astronomical object.

Fig. 3: Nebula, performed at Launch in Toronto, ON, April 2016.

The above abstraction can be calculated in under one second, and can therefore be applied to a live video sequence. The frames are interpolated to improve smoothness, and projected onto a performer, creating the effect of the performer appearing to be made of stars. This piece has been performed at two public showcases in Ontario, Canada.

Acknowledgments

Thanks to Audrey Chung for assistance on set, Angola Murdoch and Femmes du Feu for the invitation to perform Nebula at their events, and to the contemporary circus community of Ontario for their support.

References

[1] S. Greenberg, A. G. Chung, B. Chwyl, and A. Wong, "TIGGER: A texture-illumination guided global energy response model for illumination robust object saliency," in *CRV*, 2016.