# Exemplar Based Filtering of 2.5D Meshes of Faces

Leandro Dihl, Leandro Cruz and Nuno Gonçalves[1]

[1]Institute of Systems and Robotics
University of Coimbra, Portugal

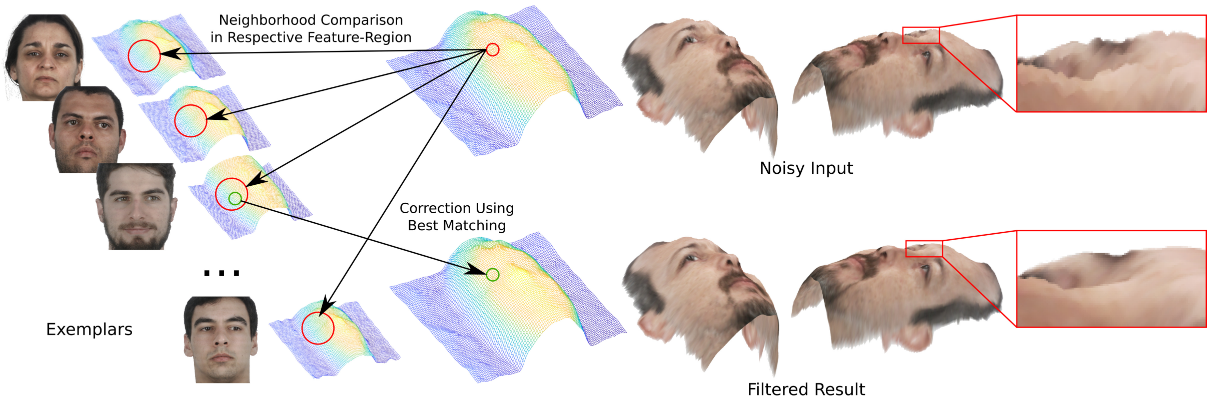


**Figure 1:** *Our content-aware filtering pipeline*

**Abstract**

*In this work, we present a content-aware filtering for 2.5D meshes of faces. We propose an exemplar-based filter that corrects each point of a given mesh through local model-exemplar neighborhood comparison. We take advantage of prior knowledge of the models (faces) to improve the comparison. We first detect facial feature points, and create the point correctors for regions of each feature, and only use the correspondent regions for correcting a point of the filtered mesh.*

**CCS Concepts**

•***Computing methodologies*** $\rightarrow$ ***Mesh geometry models;*** *Mesh models; 3D imaging; Reconstruction;*

## 1. Introduction

In the last years, it has been increasing the number of applications using faces captured from consumer 2.5D cameras. Despite of the successive evolution of these devices, the resultant mesh is still very noisy generating mistakes on person recognition, pose detection or facial expression recovery.

The input of our method is an RGBD image obtained using sensors like Kinect or Light Field cameras, providing a 2.5D regular surface - a matrix of each surface point height (height field). This structure is fundamental for our filtering step, since it allows to use a point correction approach based on texture synthesis methods [WLKT09]. This method will be presented in Section 2.

Differently than texture synthesis, our method has prior knowledge about the structure of faces. So, although the model and exemplars have different macro-features (for instance, the model face is larger or lengthier than all exemplars), one can match local features because their respective parts still have intrinsic geometric similarities. The pipeline is illustrated in Figure 1.

In this work, we present a filtering method for meshes of faces preserving intrinsic features. Particularly, our filtering is based in exemplar-based neighborhood matching. We use the geometric nature of the model to improve the matching (Section 3). Furthermore, we use facial feature points to define regions in which all points have intrinsic geometric similarities (Section 4). Then, a point on the filtered model is just compared to points in exemplars at correspondent regions. These two steps allow us to propose a content-aware method able to remove noise of a 2.5D face (by model-exemplars neighborhood comparison), as well as typical filtering methods [BPK*07]. The methods also fixes some small macro structures artifacts (by correspondent region comparison).

## 2. Model-Exemplars Neighborhood Comparison

The neighborhood of each point $M_p$ on the model is compared to the neighborhoods at the exemplars, and the position of $M_p$ is replaced by the position of the respective best matching. A neighborhood is a window containing $N \times N$ values in height field, centered at the $M_p$. The search for the most similar neighborhood on exemplars is performed by a Nearest Neighbor method, implemented using a Kd-Tree, in conjunction with PCA for dimensionality reduction [HJO*01].

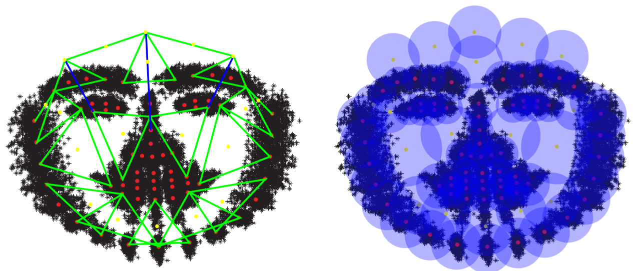

**Figure 2:** *Facial-Feature-based Model Covering [WGT*18].*

## 3. Improvement of Neighborhood Matching

Our method is constrained to a specific type of models: faces. Even so, the filtered model can contain some macro characteristic that does not exist in any exemplars. However, each $M_p$ may have micro characteristic similar to some point in the exemplar (i.e., a similar height variation into neighborhoods). In this case, we do not make a direct comparison (Euclidean distance of the neighborhoods), but we first make an adjustment to comprise these similarities

Our neighborhood comparison metric of a point $M_p$ on model and a point $E_q$ on exemplar is given by:

$$dist\,(M_p, E_q) = \sum_{i=0}^{N}\sum_{j=0}^{N}\left(M_p(i,j) - \left(E_q(i,j) - \bar{h}\right)\right)^2 \qquad (1)$$

where $\bar{h}$ is the factor that minimizes the difference among the neighborhoods. Let $d(h) = \sum_{i=0}^{N}\sum_{j=0}^{N}(M_p(i,j) - (E_q(i,j) - h))^2$ then $\bar{h} = argmin\; d(h)$. Thus, $\bar{h}$ is the solution of $d'(h) = 0$:

$$\bar{h} = -\frac{1}{N^2}\sum_{i=0}^{N}\sum_{j=0}^{N}(M_p(i,j) - E_q(i,j)) \qquad (2)$$

If the neighborhood of $\bar{q}$ is the best matching for $M_p$, then the position of $M_p$ is replaced to $E_{\bar{q}} - \bar{h}$.

## 4. Facial-Feature-based Model Covering

The major characteristic of our filtering process is to constraint the search to corresponding regions. It has two advantages: to diminish the amount of neighborhood comparisons, and to use only points that have similar micro features (reducing outliers). Due to the searching process acceleration, we can use more exemplars for correction, enhancing the robustness of the method.

The subdivision process shall meet three conditions:

1. All faces must have feature points at correspondent places;
2. The union of regions must cover the whole face;
3. The area of each region must be inversely proportional to the average of expected noise.

The first condition is achieved by using Constrained Local Models [WGT*18] for defining facial features. Because these features set does not meet the condition (2), we add some further points (Figure 2), generated by geometric operations that depend only on previously obtained features (and therefore, these points also satisfy condition (1)).

To meet the condition (3), we first perform a deformation of the models according to feature points, in such way that the respective features, after the warping in all models, are at the same place. Next, we compare each model to a filtered version (using classical filtering approaches) and create an error map (norm of difference). Following, we calculate the expected error of the region around each feature point. Then, we solve a min-max problem to find the smallest size for the larger region such that all sets of regions cover the respective model. Finally, the region size is defined to satisfy the condition (3).

## 5. Conclusions and Future Work

These first experiments were performed using single shot models from light field cameras. The exemplars set was previously improved by combining typical filtering approaches. A direction of future work is the use of better models (e.g. 3D scanners models). Moreover, we can also use partial models (accurate height fields of specific parts), since we can identify the respective feature points.

Another direction for future work is to define a validation metric for the results. It could be based on the measure of the variance of each point. Indeed, the proposed method reduces the overall variance. In addition, it is possible to use the symmetry of faces for creating this quality measure. Furthermore, we can define a metric by a comparison of the filtered model with a facial template deformed by the facial feature points.

Finally, our filtering approach is based on a division of the model in regions in which all points have an intrinsic geometric similarity. We presented how to define these regions for the specific case of faces by the use of facial features. For other types of models, it is necessary to use a feature detector that satisfies the three conditions presented on Section 4. A future work direction is to define general descriptors that can be used for general purpose filtering.

## References

[BPK*07] BOTSCH M., PAULY M., KOBBELT L., ALLIEZ P., LÃL'VY B., BISCHOFF S., ROSSL C.: Geometric modeling based on polygonal meshes. In *SIGGRAPH Course Notes* (2007). 1

[HJO*01] HERTZMANN A., JACOBS C. E., OLIVER N., CURLESS B., SALESIN D. H.: Image analogies. In *SIGGRAPH* (2001). 2

[WGT*18] WANG N., GAO X., TAO D., YANG H., LI X.: Facial feature point detection: A comprehensive survey. *Neurocomputing* (2018). 2

[WLKT09] WEI L.-Y., LEFEBVRE S., KWATRA V., TURK G.: State of the art in example-based texture synthesis. Eurographics STAR. 1