Segmentation, characterization and superimposition of PSP damage-induced dental X-ray artifacts

* Introduction
  + Background
    - Phosphor Storage Plates (PSPs) and their use in dental imaging; comparison to complementary metal oxide semiconductor (CMOS).
    - Problem of producing image artifacts that comes with using PSPs.
    - Concerns and evidence that PSP artifacts could impede diagnosis.
    - How the clinics are dealing with that issue now and the need for an accurate predictor to avoid missed diagnosis and improve economic efficiency.
    - Previous attempts of classifying and quantifying artifacts and their shortcomings.
  + Research question
    - Research question: Can we segment, characterize artifacts and superimpose them onto clean teeth images for future studies?
    - Hypothesis: We can segment, characterize and superimpose the artifacts using an algorithmic and machine learning approach.
  + Objectives
    - Segment the artifacts from the image;
    - Quantify and characterize the artifacts;
    - Cluster the segmented artifacts;
    - Superimpose artifact images onto clean teeth images to mimic real dental images taken with damaged PSPs.
* Methods
  + Preprocessing for segmentation
    - Cross-correlation analysis to remove the “C” in the image.
  + Segmentation
    - Density based spatial clustering algorithm with noise (DBSCAN) with a little description of theory behind it.
  + Clustering
    - Computation of descriptive features and K-means. (how we decided on five)
    - Reducing dimensionality of raw data then feed to clustering algorithm combined with clustering algorithms such as k-means or Gaussian mixture models.
      * CNN Autoencoder
      * Principle component analysis
  + Preprocessing for Superimposition
    - Contrast limited histogram equalization.
  + Superimposition
    - Linear combination of two images.
    - Linear combination where different weights are applied to pixels in different intensity intervals.
    - An exponential weighting function across all pixel values combined.
* Results
  + Preprocessing for segmentation
    - Show the results of C removal.
  + Segmentation
    - Show results of DBSCAN with thresholding
  + Clustering
    - Show the results of k-means clustering based on feature computation together with the computed feature summary for each cluster.
    - Show the results of the autoencoder for dimensionality reduction.
    - Show the results of clustering and the descriptive statistics of them
  + Preprocessing for superimposition
    - Show results of contrast limited histogram equalization with a comparison to the original artifact image: subtle artifacts are now better shown on the image.
    - Show image histogram before and after the equalization.
  + Superimposition
    - Show results of whole image linear combination using original artifact images and equalized artifact images (side by side with images taken with the damaged plate as a control). Discuss the problems:
      * While using original images, subtle artifacts are not obvious enough in the final superimposed images.
      * While using equalized artifacts, artifacts with high intensities are not bright enough in the final superimposed images.
    - Show results of using different weights for pixels within different intensity regions (using original and equalized artifacts in comparison to controls).
      * Limited improvement if weights for neighboring thresholds are similar.
      * Visible inconsistency and gaps of intensity within regions that contain pixels from multiple thresholded intervals if weights for neighboring thresholds are distinct.
    - Show results of using exponential weighting function (using original and equalized artifacts in comparison to controls).
      * Show how the parameters for the exponential function were determined.
      * Show the final result: superimposition done with equalized artifact images, combined with the incorporation of the bright artifacts from original artifact images.
* Discussion
  + Briefly mention the Canny edge detector and its success in picking up line/crack artifacts
  + Based on the results from DBSCAN, discuss the suitability and limitations of this algorithm. State the need for thresholding and explain the choice of parameters and the threshold.
  + Discuss feature computation and the rationale for the features chosen. Based on the features for each determined cluster, discuss the results of k-means clustering and the loss of information due to feature computation.
  + Explain the curse of dimensionality and the theory behind an autoencoder. Discuss Gaussian mixture model and a theoretical estimation of maximum features the model can take in.
  + Discuss PCA for dimension reduction.
  + Discuss the clustering results and the descriptive statistics of them.
  + Discuss the histogram equalization and the different characteristics of the original and equalized artifact images.
  + Discuss the three superimposition algorithms used, their limitations and how they have improved from their precursor.
  + Discuss the use of original artifact images and equalized artifact images, compare the results they produce. Discuss the reasons for the final approach taken, which used the equalized images for the bulk part and the original images for the bright artifacts.
  + Discuss the need for a validation process for superimposition, our attempts of SSIM and SIFT. Point out the difficulties of such validation on our dataset of superimposed vs real ones, which include the shift of frame of the teeth in the back ground, the differences in scale between images taken with CMOS and PSPs, and the insufficient consideration of artifacts comparing to teeth background by the similarity metrics.
* Future directions
  + Hypothesize the impact of segmented artifacts based on their post-superimposition location on the teeth images.
  + Use the superimposed images to obtain golden standards for whether an artifact will impede diagnosis.
  + Following that, supervised machine learning can be applied to investigate the relationship between the artifacts (now segmented by DBSCAN) and their ability to impede diagnosis.
  + Ultimately the goal would be to establish a model that can accurately predict whether a given PSP will affect diagnosis, and therefore, should be discarded.
* References