Vedant Puri
veantpu@andrew.cmu.edu

Homework 3

Question 1: Contrast enhancement, filtering, segmentation, and registration are four important image processing techniques. For each of them, please describe what it is used for.

Answer: We describe the following techniques in image processing:

1. **Contrast Enhancement:** Contrast enhancement is an image processing technique to exaggerate visible differences between adjacent structures in images. Contrast enhancement is achieved by increasing the difference between "dark" and "light" parts in an image, or touching up various colors. It is usually a pre-processing step aimed at improving perception of different objects by the human eye, or to ease downstream classification in downstream algorithms. To adjust the brightness and enhance the contrast, we can use a stretching function or a windowing function and apply it pixel-wise to the image. It is adaptive, multi-scale, and anisotropic. In adaptive contrast enhancement, a new intensity is assigned to each pixel according to an adaptive transfer function based on local statistics.

2. **Filtering:** Data from the real-world is often riddled with noise that needs to be *filtered out*. In image-processing, this is a necessary pre-processing step as many algorithms cannot work with noisy data. While contrast enhancement is a pointwise (pixel-wise) operation, filtering necessarily considers multiple pixels (as frequency information is necessarily global or over an element), and is applied via a Fourier transform or a convolution. In convolutions, the filter kernel is convolved with the image to smooth noisy data, or gain local information about the image (eg. derivative). The Fourier transform, on the other hand, is a global operation that decomposes a signal into sinusoidal waves of varying frequencies. By the convolution theorem, applying a convolution is equivalent to applying a pointwise operation in the frequency space. Critically, convolutions and Fourier transforms are translation invariant.

3. **Segmentation:** Segmentation or classification is a process to partition an image into multiple regions or sets of pixels. It is used to locate objects and boundaries in images. Segmentation in medical imaging is used to locate tumors, measure tissue volumes, assist with computer guided surgeries, and study anatomical structure. Image segmentation is an open problem most techniques are domain specific A major task in segmentation is edge detection that aims to locate boundaries between objects in images which utilizes filtering and contrast enhancement as pre-processing steps.

   The watershed transform is a strategy for image segmentation. The term 'watershed' refers to a ridge that divides areas drained by different river systems, and a catchment basin is the geographical area draining into a river. When we think of images as describing surfaces $\phi(x, y)$, with bright areas as high and dark as low, it is natural to think in terms of catchment basins and watershed lines. Pixels with highest gradient magnitude intensities (GMIs) correspond to watershed lines, which represent region boundaries. Water inside a region flows downhill to a common local intensity minima (LIM). Pixels draining to a common LIM form a watershed region. The resulting watershed can be hierarchically organized into a graph with tree-like structure. The graph hierarchy can be used to navigate the image with interactive, dynamic segmentation.

4. **Registration:** Registration is the process of finding spatial transforms that map points from one image to corresponding points on another image. Medical image registration has many applications. Repeated image acquisition from a subject is often used to obtain time series information that captures disease development, treatment progress, and contrast bolus propagation. Additionally, correlating information from different imaging modalities requires some form of registration. Image registration determines geometric transformations that map common structures in images to each other. A common methodology in registration is to map structures in images to a common latent space and then do mapping between images via the common latent space.

Question 2: Select one image segmentation paper and one image registration paper from the internet (different from the references used in the lecture notes), explain the main algorithms and discuss the pros and cons. You don't need to submit the pdf file of your selected papers, but please list the reference/citation information.

Answer:

1. **Image Segmentation:** The recent trend is to use deep convolution neural networks and image transformers for image segmentation. Convolutional Neural Networks with contracting and expanding paths have shown prominence for the majority of medical image segmentation applications since the past decade[2]. In CNNs, the encoder plays an integral role by learning both global and local features and contextual representations. Despite their success, the locality of CNN layers limit its capability of learning long range spatial dependencies. Transformer neural networks, inspired by techniques in natural language processing are becoming increasingly popular as they can learn sequence representations with long-range, global, multi-scale information.

2. **Image Registration:** We take a look at the problem of shape registration on 3D meshes of obtained from biomedical data in [1]. Spatial registration methods usually require well-defined features or landmarks to map shapes between images which can be very hard. Spectral methods, on the other hand, do not need any landmarks. Shape spectrum, inspired by the Fourier transform in signal processing, considers discrete meshes as graphs and looks at the spectrum of the its graph Laplacian and other operators.

   [5] describes the field of spectral mesh processing where of eigenvalues, eigenvectors, or eigenspace projections derived from appropriately defined mesh operators are used to carry out desired tasks. Spectral analysis of mesh geometry based on a graph Laplacian has been studied since 1995, with many new methods coming up in recent decades. The new trend is to use spectral geometry processing ideas as layers in deep neural networks. This allows machine learning models to learn on mesh invariant features, and generalize faster[4].

References

[1] Hajar Hamidian. *3d Surface Registration Using Geometric Spectrum Of Shapes*. Wayne State University, 2019.

[2] Ali Hatamizadeh et al. "Unetr: Transformers for 3d medical image segmentation". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2022, pp. 574–584.

[3] KKD Ramesh et al. "A review of medical image segmentation algorithms". In: *EAI Endorsed Transactions on Pervasive Health and Technology* 7.27 (2021), e6–e6.

[4] Nicholas Sharp et al. "DiffusionNet: Discretization Agnostic Learning on Surfaces". In: *ACM Trans. Graph.* (2022).

[5] Hao Zhang, Oliver Van Kaick, and Ramsay Dyer. "Spectral mesh processing". In: *Computer graphics forum*. Vol. 29. 6. Wiley Online Library. 2010, pp. 1865–1894.