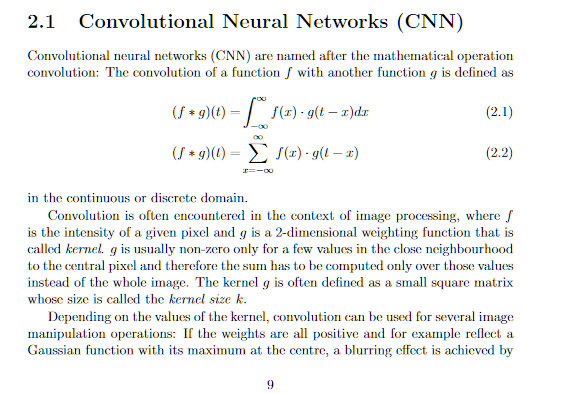
**2. Theory (2-3 pages)**

A simple way of determining the relationship between two separate datasets is by comparison of patterns with template data. Particularly, by a method of single or multi element template matching via correlation calculation. Correlation, also covariance, is a measure of the strength of a linear relationship between two or more quantitative variables. So, given two independent variables, that both vary in a way such that they are both related to a third in the same way such that their behaviour is positively correlated as a result of their individual relationship with the third variable. Thus, it is useful to derive this relationship for data that is to be compared. This relationship is quantified as the cross correlation between sets of data. Computationally, two classical approaches to find the cross-correlation, spatial and spectral, are implemented in this project.

1. Pattern matching[[1]](#footnote-1)
   1. Cross correlation

Statically, cross correlation provides a measure of association between datasets. The correlation value may be determined from a normalised unnormalized weighting of the data set. Quantitatively it provides the degree of strength of a linear relationship between two or more quantitative variables. The value of the cross correlation is determined in general by eq1 and may be generalized for higher dimensions.

*eq1*

Where f1 and f2 are the two different data sets, n is a discrete element within the data, is the shifted position, and c(s) is the correlation coefficient. The position where the coefficient is largest or smallest, s, identifies the shift position for which one data set most greatly matches the other. The effect of the offset position is visualised below (figure 1). The position of greatest correlation appears as a peak in the graph of correlation coefficient plotted against shift position (figure 2).

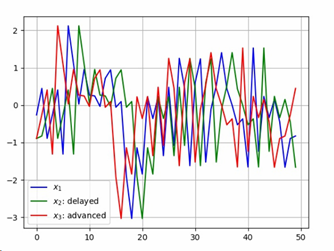


Figure 1. Code to generate in appendix A

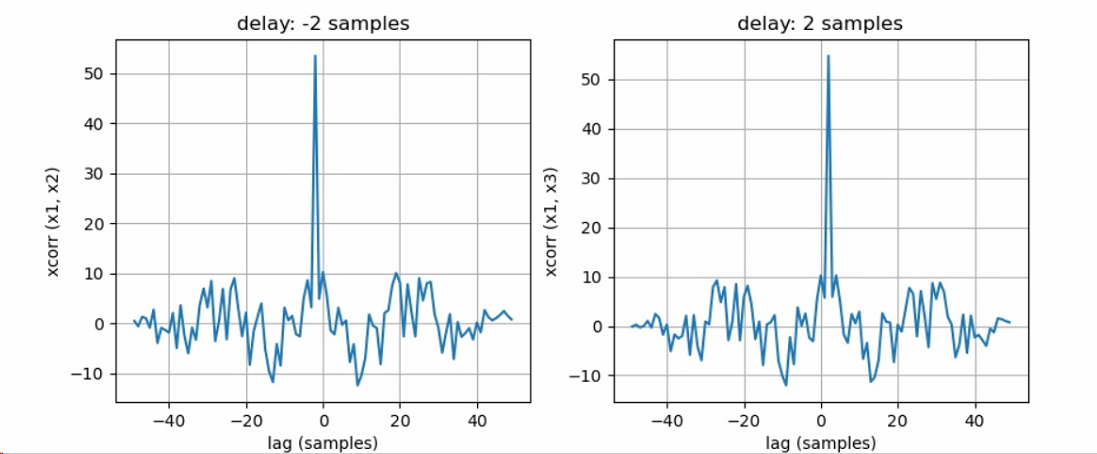


Figure 2

* 1. Normalised cross correlation[[2]](#footnote-2)[[3]](#footnote-3)

Direct correlation, unless autocorrelation, is dependent on the amplitude of the data compared. This dependency is eliminated by normalising the correlation coefficient, referred to as the normalized cross-correlation. Normalized cross correlation is used to evaluate the degree of similarity between any two data sets. It has the advantage of less sensitivity to linear changes in the amplitude of the quantitative variables under comparison. [[4]](#footnote-4)

The normalised cross correlation is divided by the product of the autocorrelation magnitude as this is the largest that correlation value the data may have. The autocorrelation of discrete-time signals is a useful mathematical concept that helps measure the maximum or total energy of a signal. Calculation results have a fixed range 0 to 1. If there is a large difference in amplitude between the matched regions, then the modified normalization potencies the correlation to zero (i.e. effectively uncorrelated). Finally, contrast normalization allows matching with regions that are principally uniform except for random changes or texture variations.

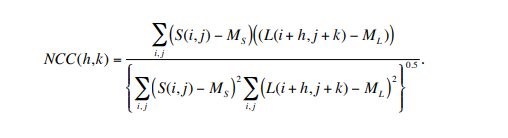
* 1. Spatial

Spatial cross-correlation is a measure of the strength of the relationship between two temporal data sets. It is found by convolving each data set across each other that may expose patterns in the input data. This method is useful for finding the overlapping pattern over a periodic signal covered by noise. Specifically, this is implemented for two signal files that are analysed in this section. Since the two signals compared are given to be different, cross correlation is implemented. The cross correlation will reveal how much one signal has been shifted, for the piecewise continuous case of two signals with appropriately short sampling, to be congruous with another comparative signal. Defined formally:

*eq1*

The similarity of the two signals is given thus by the correlation coefficient *c* for a given lag *s* between the two signals. [[5]](#footnote-5)

Further, this is made more general by the normalized cross-correlation in 2D. The normalized cross-correlation forces uniformity across the data sets where there is a wide spread of intensities/amplitudes for the data.

* 

The Normalized Cross Correlation is confined to a range from the inclusion of the denominator so that the resulting correlation ranges from -1 to 1. A perfect match has a value of 1. ML ≡ ML(h,k) is the mean of the subsection of the large image at offset (h,k), N is the number of pixels in the small image and NCC(h,k) is the normalized cross correlation metric at offset (h,k). The numerator is essentially a simple cross correlation but using a zero mean small image and zero mean subsections of the larger image. The mean subtraction mitigates amplitude differences between the two images. The denominator is included so that the resulting correlation metric ranges from -1 to 1.

In spatial cross correlation, a weighting is needed to be derived at each alignment. This is computationally intensive as the time complexity for this process is N2 where N is the input size of the larger array. The process of cross-correlation can be performed more efficiently for large inputs by transforming the inputs into the frequency domain.

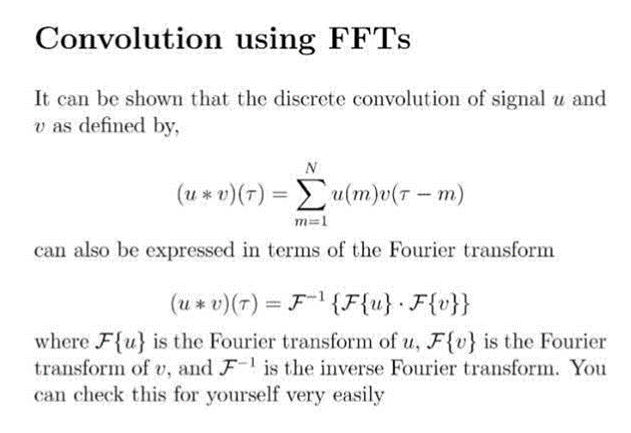
* 1. Spectral

The Fourier transform produces another representation of a signal, specifically a representation as a weighted sum of complex exponential terms of the frequency spectrum of the data. Spectral cross correlation utilizes the same principles from which spatial cross correlation is are based, however operations between two data sets are made only in frequency space.

A Fourier transform is a signal processing technique which decomposes a function on the time domain into a function in the frequency domain. Times series, spatial data, is converted into frequency data using the Fourier transform. Fourier transforms decompose the frequencies of periodically sampled data and map them to corresponding amplitudes in the frequency domain. An approach not using normalization and directly computes the simple cross correlation, C(i,j), using forward and inverse Fourier transforms. A basic principle of Fourier transforms is that convolution in the spatial domain is equivalent to multiplication in the frequency domain. Likewise, correlation in the spatial domain is equivalent to multiplication in the frequency domain using the complex conjugate of one of the transformed arrays. Thus, either convolution or cross correlation can be used to perform spectral cross correlation. The cross correlation of f(t) and g(t) will have the same output as the convolution of f(t) and g(t) but reversed. This results in the formula below:

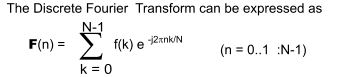
Xxxxxxxxx eqn

As such the discrete convolution of two series signals may be determined by finding the inverse of product of the Fourier transforms of both signals. [[6]](#footnote-6)



This relies on the property that multiplication in the frequency domain is equal to convolution in the spatial domain. Because of this it is easier to apply multiple filters as products in the frequency domain, where they would have been successive convolutions in the time domain.

Since time series data is sample discretely, a summation analogue of the continuous Fourier transform is used:



In practice the discrete Fourier transform (DFT) is implemented typically using the Fast Fourier Transform (FFT).

In contrast to spatial correlation, the normalised cross correlation does not have a minimal frequency domain expression. It cannot be directly computed using the more efficient FFT (Fast Fourier Transform) in the spectral domain. Its computation time increases dramatically as the window size of the template gets corpulent.

1. Stereo vision
   1. Depth mapping[[7]](#footnote-7)

Stereovision aims to reproduce the depth determination capability of human vision. A depth map records the distance information with respect to a viewpoint in a 2D image of digitised real-world coordinates. The proposed algorithm uses two images of the same scene. Pixel to pixel matching is used between two images, left and right, to find the corresponding point in the same scene from different views. As such, the same point is identified in 3D from the 2D image information. The left and right images are assumed to be taken with two digital cameras. They capture images of the same scene, at the same time. Given some horizontal displacement between the cameras, which may be approximated to that the spacing between two eyes. Given the geometric description of the location of the camera, separations and pixel location, a depth map may be created using a method binocular disparity.

The disparity value of a point is often interpreted as the inverse distances to the observed objects. Thus, disparity is inversely proportional to Depth. Therefore, finding the disparity is essential for the construction of the depth map. Darker regions in the Depth-Map are created for signifying that an object is far away, and this darker colour gradually decreases to brighter with decrease in depth and finally becomes white for closer objects.

­

1. Practical considerations
   1. Data set size

Along with physical computational power, the efficiency of a cross correlation algorithm will vary with size of the data being compared. In the case of image data, which is the primary focus of the project, the degree of computational complexity will depend upon image size and resolution. The nature of image data is such that the size of the data set, a digital image a[m, n], described in a 2D discrete space is derived from an analogue image a(x, y), in a 2D continuous space through a sampling process. The process is often referred to as digitisation. The digitisation relates image data to pixel location. The data sets used in this study consider one- and two-dimensional analogue information that is already pre-digitised in text or image-based formats.

* 1. Time complexity[[8]](#footnote-8)

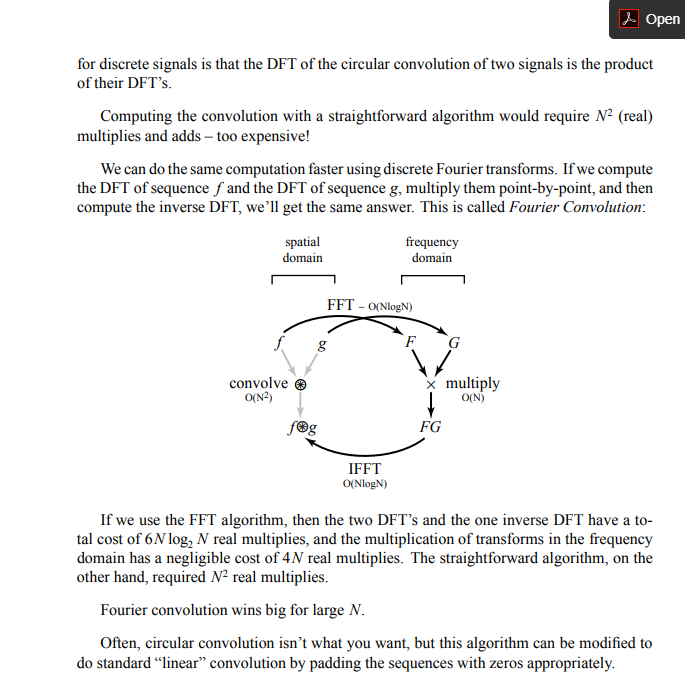
Direct calculation of NCC is computationally intensive and the time cost is proportional to the region size. Assuming the average size of the matching regions is R, the computational complexity to match two images with image size of M and disparity range of D is O(M ×R× D). Since M, R and D are usually large. Since stereo matching usually needs to match each pixel across a disparity space for a whole image, the computational efficiency of the pixelwise similarity is necessary for a fast depth determination.

* + 1. Spatial

The spatial cross-correlation methods considered in this project utilises convolution to relate the elements of the two compared data sets. The computational complexity for a N×M convolution kernel implemented in the spatial domain on an image of N × M is O(N2) where the complexity is measured per pixel on the basis of the number of multiplies-and-adds(reference).

* + 1. Spectral

The Fast Fourier Transform, which is first introduced by Cooley and Tukey, 1965, the Cooley-Tukey algorithms. The time upper bound time complexity and operation counts is O(N log(N)). Thus the FFT is an optimized algorithm that reduces the time complexity of the DFT from N^4 to NlogN. [[9]](#footnote-9) The computational complexity per pixel of the Direct Fourier approach for an image of N2 × N2 and for a convolution kernel of N × N is O(log N) complex MADDs independent of N.



1. <https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1045&context=kin_pubs>

   <https://lib.dr.iastate.edu/kin_pubs/46> [↑](#footnote-ref-1)
2. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.675.1379&rep=rep1&type=pdf [↑](#footnote-ref-2)
3. <https://www-scientific-net.eu1.proxy.openathens.net/AMR.860-863.2800.pdf> [↑](#footnote-ref-3)
4. https://imagemagick.org/docs/AcceleratedTemplateMatchingUsingLocalStatisticsAndFourierTransforms.pdf [↑](#footnote-ref-4)
5. Goodman, Fourier optics [↑](#footnote-ref-5)
6. Proof in appendix [↑](#footnote-ref-6)
7. https://projet.liris.cnrs.fr/imagine/pub/proceedings/ICIP-2009/pdfs/0002357.pdf [↑](#footnote-ref-7)
8. <http://hit.skku.edu/wp/wp-content/uploads/1-s2.0-S1051200410000047-main.pdf>

   https://projet.liris.cnrs.fr/imagine/pub/proceedings/ICIP-2009/pdfs/0002357.pdf [↑](#footnote-ref-8)
9. Reference [↑](#footnote-ref-9)