Chapter 4
Effect of Imaging Conditions and Image Quality on Image Registration

4.1 Introduction

Image registration forms a basis for a wide variety of applications in Computer Vision, Medical Imaging, Remote Sensing, and Satellite Communication etc. The methods used for image registration are generally divided into two categories: 1) Extrinsic Methods: These methods are based on some external objects placed in the scene. 2) Intrinsic Methods: These methods are based on image information accessed in form of pixel intensity values, color information etc. to perform further formulations with respect to the requirements of a particular application [Wyawahare et al. 2009]. Therefore, while designing an Augmented Reality (AR) application, marker based or markerless, great emphasis is laid upon the features to be extracted from the image scene [Ferrari et al. 2001]. Stability of these extracted features define the correct estimation of position of virtual objects that are to be integrated with the view of real environment. However, use of markers for this purpose have only limited applicability. Therefore, for incorporating AR in a wide variety of applications, there is a need for detecting affine invariant and stable natural features from an image [Chen et al. 2010].

This chapter presents a comparative analysis of six feature detectors, which could be used for the image registration process in AR. The study is divided in two setups where the 1st setup involves the behavior analysis of three feature detectors (Scale Invariant Feature Transform (SIFT) [Lowe 2004], Affine-SIFT (ASIFT) [Yu and Morel 2011] and Speeded Up Robust Features (SURF) [Bay et al. 2008]) with respect to illumination and blur conditions in an image and its respective quality. The dataset [Appendix A.1] used for this study contains seven image-sets with image variants that determines different illumination and blur conditions of medial, natural and structured scenes. The quality quantification is done using four No-Reference Image Quality Assessment (NR-IQA) metrics (Spatial and Spectral entropies based Image Quality Assessment (SSEQ) [Liu et al. 2014], Naturalness Image Quality Evaluator (NIQE) [Mittal et al. 2013], Blind/Reference-less Image Spatial Quality Evaluator (BRISQUE) [Mittal et al. 2012] and BLind Image Integrity Notator using Discrete Cosine Transform (DCT) Statistics-

II Index (BLIINDS-II) [Saad et al. 2012]) metrics and four Full-Reference Image Quality Assessment (FR-IQA) metrics (Structure SIMilarity Index (SSIM) [Wang et al. 2004], Multi-Scale Structure SIMilarity Index (MS-SSIM) [Wang et al. 2003], Mean Square Error (MSE) [Avcibas et al. 2002], and Normalized Cross-Correlation (NK) [Eskicioglu and Fisher 1995]). For 2nd setup, few observations from 1st Setup are taken into consideration to extend the study for six feature detectors (Harris-Affine [Mikolajczyk and Schmid 2002, Mikolajczyk and Schmid 2004], Hessian-Affine [Mikolajczyk and Schmid 2002, Mikolajczyk and Schmid 2004], Maximally Stable Extremal Regions (MSER) [Matas et al. 2004], SIFT, ASIFT and SURF). These feature detectors are tested on 48 images [Appendix A.1] (divided into eight image-sets with six images in each set) varying in terms of change of viewpoint, scale, blur, illumination and JPEG compression ratio. Quality Quantification in this case is done using two NR-IQA metrics (SSEQ and BRISQUE) and two FR-IQA metrics (SSIM and MS-SSIM). The novelty of the this study is the usage of image quality metrics and Pearson Coefficient to study the traits of feature detectors in terms of number of detected keypoints in an image and number of matches found between two images with respect to image quality.

4.2 Imaging Conditions

Challenges faced by Augmented Reality Systems due to varying Imaging Conditions: Various changes in imaging conditions like, translation, scaling, rotation, illumination, blur and affine as discussed in Chapter 3 [Section 3.3], may cause incorrect estimation of position of virtual objects as image registration becomes challenging in such scenarios. Moreover, views of real scene captured as images may also contain different distortions where some may have been originated from the optic characteristics of image sensors and occurrence of others may be described due to some specific scene setups or objects.

In any form, unfavorable imaging conditions makes the receptivity of an AR system to degrade. Therefore, in general, some reasonable assumptions defining a scenario for a particular application could be kept in mind while designing an AR system. For example, in case of indoor or built environment where objects are a part of a controlled surrounding, virtual object position estimation, once accurately computed, could be contained by managing the initial environment parameters. But, such a task becomes challenging when AR system is designed for an outdoor environment where imaging conditions could drastically change due to many factors as the state of the environment being captured greatly affects the augmented display [Lin et al. 2006, Lin et al. 2009]. Less preferable conditions, in terms of nature of the environment being captured, impede the understanding of the world around us. The main source of obstruction in both indoor and outdoor environments is the illumination condition. Lighting affects the actual scene content in form of incorrect display of color in images, resulting in incorrect augmentation

[Li et al. 2012]. Also, highly varying lighting can make projection difficult as very bright environments can limit projection. For outdoor scenes, two factors affecting light intensity are: time of day and weather (for example, clouds, fog, and rain can limit visibility), interfering the visibility of objects at that time. As in indoor conditions, strong light (both natural and artificial) can cause reflections and lens flare.

The added virtual information to a scene is often attained by either overlaying the virtual information to real world objects (e.g., replacing the Two Dimensional (2D) book image of an atom by a virtual Three Dimensional (3D) model), or by adding the virtual content to the real scene (for example, a label defining a particular object). The user is therefore expected to be able to distinguish both kinds rightly. Main barriers to perceptually correct augmentation could be also defined by image depth estimation factors that are usually interrelated. However, incorrect depth estimation of a scene is the most common perceptual problem in AR applications, hindering the spatial relationship between the users perspective, the objects in view, and the overlaid/inlayed virtual information.

4.3 Effect of Imaging Conditions on Image Registration

Correlation of variations in Imaging Conditions with performance of Image Registration methods: Image registration methods rely on extracted features from an image scene which are processed in the form of interest points. These interest points should have a well-defined and distinctive neighborhood or a well-textured area in the environment which makes them perceptible and easy to retain in a series of image sequence for achieving an accurate estimation of the position of virtual objects in the real environment while developing an AR system. But, major issue with interest points is that, in some scenarios, identifying the distinctiveness of the extracted interest point becomes difficult [Baumberg 2000]. For example: in indoor scenes, processing localization techniques based on natural features becomes challenging where blank walls usually occur, making the task of defining a feature distinctively hard enough. So, if we assume that feature descriptors that process on natural features are usually designed to be illumination invariant, this assumption can only be true for scenes where environment contains actual physical features, e.g., outdoor environments, where natural features could be distinctively defined with the help of various characteristics, e.g., intensity value, color etc. But at times, extracted natural features from an outdoor environment does not relate to real physical features, i.e., blobs are often formed as a result of shadows cast by objects in the scene, corners and lines occur and dynamically move as the lighting or weather conditions change. As a result, an overwhelming number of outliers and mismatches effect keypoint matching accuracy, irrespective of the choice of matching algorithm [Lin et al. 2006].

Image registration is also affected by the kind of transformation an object undergoes in two or more consecutive image frames. Depending upon the transformation model, parameters for registration changes as per the requirements listed in [Behzadan et al. 2015]. For example, in a rigid transformation model, which preserves relative distances of points, translation and rotation parameters are estimated, whereas in affine transformation model, which may not preserve collinearity, non-rigid transformation parameters are estimated. For a more intricate transformation model, i.e. perspective projection, affine transformation parameters along with the transformations of panning and tilting are taken into account.

4.4 Image Quality Metrics & Image Quality Variations

4.4.1 Pearson Coefficient

In this research, to understand the correlation between every pair of image in each image-set, Pearson Coefficient (correlation coefficient) is used. Pearson coefficient shows the linear relationship between two sets of data. Its values range from 1 (for two images whose intensities are perfectly linearly correlated) to −1 (for two images whose intensities are inversely correlated to one another). Values near zero reflect distributions of probes that are uncorrelated with one another. For more details on this topic, please refer to Section 3.2 in Chapter 3.

4.4.2 No-Reference Image Quality Assessment

Introduction: Referring to the detailed explanation of NR-IQA methods done in Chapter 3 [Section 3.4.1], NR-IQA involves quality evaluation of an image based on only test image. In present research, four NR-IQA metrics (SSEQ, NIQE, BRISQUE and BLIINDS-II) are used for quality assessment of images.

Effectiveness of using these metrics.

SSEQ: Two-stage framework of SSEQ is initiated by distortion classification followed by quality assessment. It incorporates local spatial and spectral entropy features of distorted images for understanding the distortion type. [Liu et al. 2014]. The method has been statistically proven to be superior to many other NR-IQA metrics, example: Blind Image Quality Index (BIQI) [Moorthy and Bovik 2010], Distortion Identification-based Image Verity and Integrity Evaluation (DIIVINE) [Moorthy and Bovik 2011] etc.

NIQE: It first constructs 'quality-aware' collection of features computed as per the Natural Scene Statistics (NSS) model. The quality of the distorted image is expressed as a simple distance metric between the model statistics and those of the distorted image [Mittal et al. 2013] Comparative study conducted by Mittal et al. [Mittal et al. 2013] shows that NIQE competes well with some of the best performing NR-IQA techniques that requires training on large databases of human opinions of image distortion. The makers of this model conclude that they have succeeded in creating a first kind of NR-IQA model that assesses image quality without the knowledge of anticipated distortions or human opinions of them.

BRISQUE: BRISQUE uses NSS of locally normalized luminance coefficients to quantify possible losses of 'naturalness' in the image. BRISQUE is computationally less expensive than other blind image quality assessment algorithms because it does not require to transform the image in other domains, making it well suited for real time applications [Mittal et al. 2012]. For wide range of transformations, BRISQUE is proven to be statistical better than some of the FR-IQA methods such as SSIM.

BLIINDS-II: Given certain extracted features based on NSS model of image DCT coefficients, the BLIINDS-II approach uses Bayesian inference model to predict image quality score. Some features that are indicative of perceptual quality are then formed by using estimated parameters of the model. Hence, BLIIND-II adopts a simple probabilistic model for score prediction and requires minimum training. Given the extracted features from a distorted test image, the quality score that maximizes the probability of the empirically determined inference model is chosen as the predicted quality score of that image [Saad et al. 2012].

Usage & Interpretation of these metrics in this research: In this research, SSEQ, NIQE, BRISQUE and BLIINDS-II, NR-IQA metrics are used for determining individual quality of images. The NR-IQA value interpretation of an image is also used for determining the best quality image in a particular image-set, which could be further used as a reference image for FR-IQA and image matching tasks. Also, keypoint detection performance behavior for six feature detectors (Harris-Affine, Hessian-Affine, MSER, SIFT, ASIFT and SURF) is reasoned on the basis of NR-IQA value of the image.

4.4.3 Full-Reference Image Quality Assessment

Introduction: As discussed in Chapter 3 [Section 3.4.2], FR-IQA metrics involve a reference image that is considered to be of an acceptable quality, and hence, the quality quantification of the deformed image is done with respect to this reference image. Present research involves four FR-IQA metrics, SSIM, MS-SSIM, MSE and NK for quality evaluation of images.

Effectiveness of using these metrics: SSIM and MS-SSIM algorithms make use of separated comparisons of local luminance, contrast and structure between a distorted image and its reference image. SSIM provides very accurate results in terms of the correlation between the quality predictions for two images (reference and the test image) and the subjective score. Since the perceived quality of an image heavily depends upon the scale at which the image is analyzed, MS-SSIM is exploited at multiple scales of an image, considering the effects of varied viewing distances [Wang et al. 2003]. MSE, on the other hand, measures the average of the squares of the errors or deviations between the reference and the test image and NK measures the similarity of two images as a function of the displacement of test image relative to the reference image.

Usage & Interpretation of these metrics in this research: In this research, four FR-IQA metrics are used for determining the correlation between one reference image in the image-set to other remaining five images. Reference image in a particular image-set is chosen using the NR-IQA value for a respective image. For example, in Graffiti image-set, if NR-IQA value for Image1 depicts its best quality in the image-set, then Image1 is taken as the reference image and the other five images are considered as test images. This routine is also followed when performing image matching i.e. as per the above example, Image1 is taken as the reference image when finding correspondences between images. Also, image matching performance for the feature detectors is analyzed and reasoned with respect to FR-IQA metric values between the image pair.

4.4.4 Image Quality due to Compression

Introduction: As it is believed that human visual system does not require all bits of luminance information that are present in the undistorted image, therefore, it seems acceptable to reduce the number of bits per pixel. However, a too large reduction may lead to a visible loss of image quality [Chapter 3, Section 3.5]. In this research, Lossy image compression, JPEG compression, is considered in one of the image-sets used for experimentation. It is treated as one of the imaging condition or image deformation present in the images and the effect of image compression on feature detectors performance is relatively analyzed.

4.5 Methodology & Experimental Setup

4.5.1 Methodology

Methodology and how to compare performance: The experiments are carried out in two setups:

1st Setup: Effect of illumination and blur change is studied for three feature detectors namely, SIFT, ASIFT and SURF [Chapter 3, Section 3.1.1] in correlation with the quality of images. Four NR-IQA metrics (SSEQ, NIQE, BRISQUE and BLIINDS-II) [Chapter 3, Section 3.4.1] and four FR-IQA metrics (SSIM, MS-SSIM, MSE and NK) [Chapter 3, Section 3.4.2] are used for quality quantification of images.

2nd Setup: Performance of all the six feature detectors (Harris-Affine, Hessian-Affine, MSER, SIFT, ASIFT and SURF) [Chapter 3, Section 3.1.1] against various imaging conditions [Chapter 3, Section 3.3] is compared and analyzed with respect to varied image quality. Quality quantification is done using two NR-IQA metrics (SSEQ and BRISQUE) [Chapter 3, Section 3.4.1] and four FR-IQA metrics (SSIM and MS-SSIM) [Chapter 3, Section 3.4.2]. Along with quality quantification of images, Pearson coefficient is used here to understand and reason the behavior of feature detectors.

Which statistics/metrics to use and how: Table 4.1 describes the corresponding tables and figures listing in the chapter with respect to 1^{st} and 2^{nd} Setup. Table 4.2 specifies the metrics used for reasoning the performance of feature detectors in terms of Keypoint Detection and Image Matching.

Table 4.1. Tables & Figures Representing Respective Performance Evaluation

<u>1st Setup</u>

	Table / Figure	**Comments**
Dataset	Figure 4.1	Seven Image-sets containing images with different Illumination and Blur conditions.
Keypoint Detection	Table 4.3	Table 4.3 presents SSEQ, NIQE, BRISQUE and BLIIND-II, NR-IQA metrics value for each image in an image-set along with keypoint detection results for three feature detectors.
Image Matching	Table 4.4	Table 4.4 presents SSIM, MS-SSIM, MSE and NK, FR-IQA metric values for corresponding pair of image in each image-set along with feature matching results.

<u>2nd Setup</u>

	Table / Figure	**Comments**
Dataset	Figure 4.2	Eight Image-sets containing images with five different imaging conditions.
Pearson Coefficient	Table 4.5 Table 4.9 Figure 4.3	Table 4.5 shows the numerical statistics for Pearson coefficient between every pair of image in each image-set and Figure 4.3 represents the corresponding graphical statistics. Table 4.9 presents the pearson coefficint for SSIM and MS-SSIM values for all eight image-sets

	Table / Figure	Comments
Keypoint Detection	Table 4.6 Figure 4.4 Figure 4.5	Table 4.6 presents SSEQ and BRISQUE, NR-IQA metrics value for each image in an image-set along with keypoint detection results for six feature detectors. Result analysis of Table 4.6 is done in Table 4.7
Image matching	Table 4.8 Figure 4.6 Figure 4.7	Table 4.8 presents SSIM and MS-SSIM, FR-IQA metric values for corresponding pair of image in each image-set along with feature matching results.

Table 4.2. Statistics/Metrics to Use and How

1st Setup

	Metric Used	Method
Keypoint Detection	NR-IQA	SSEQ, NIQE, BRISQUE and BLIINDS-II
Image matching	FR-IQA	SSIM, MS-SSIM, MSE, NK

2nd Setup

	Metric Used	Method
Keypoint Detection	NR-IQA	SSEQ and BRISQUE + Pearson Coefficient (to correlate the relationship between image pairs)
Image matching	FR-IQA	SSIM, MS-SSIM + Pearson Coefficient (to correlate the relationship SSIM and MS-SSIM values for all image pairs in a particular image-set.)

4.5.2 Experimental Setup

Language, Software and Tools used for implementation with system specification: The experiments are carried out using single threaded code on a computer with 16GB RAM and Intel® Core ™ i5-3470 CPU@3.20ghz × 4 processor with cache size of 6144 KB.

Implementation details: In 1st Setup three feature detectors, SIFT, ASIFT and SURF are studied and compared with respect to the number of keypoints detected in images and number of matches found between two images. Number of keypoints detected in images is correlated with four NR-IQA metrics values (SSEQ, NIQE, BRISQUE and BLIINDS-II) and number of matches between two images in correlated with four FR-IQA metrics values (SSIM, MS-SSIM, MSE and NK). Observations from this comparative study are used for extending the behavior analysis of six feature detectors under varying imaging conditions and image quality in 2nd Setup. All four NR-IQA and FR-IQA metrics are implemented using MATLAB implementation [NR-IQA: Appendix B.1, FR-IQA: Appendix B.2]. For feature detectors, Yu and Morel [Yu and Morel 2011] reference source code [Appendix B.3] is used for ASIFT, demo source provided by Lowe [Lowe 2004, Appendix B.3] is used for SIFT and author's binaries [Bay et al. 2008] are used for SURF [Appendix B.3].

In 2nd Setup, the six feature detectors, Harris-Affine, Hessian-Affine, MSER, SIFT, ASIFT and SURF are studied and compared with respect to the number of keypoints detected in images. These results are correlated with two NR-IQA metrics (SSEQ and BRISQUE) and Pearson Coefficient observations. The reason for not choosing NIQE and BLIINDS-II NR-IQA metrics is briefed in Section 4.6.1. The three added feature detectors in this setup (Harris-Affine, Hessian-Affine and MSER) are implemented using binaries provided by the respective authors [Appendix B.3].

For feature matching, only SIFT, ASIFT and SURF feature detectors are used in this setup for the comparative study. The reason behind this approach is that, these three detectors have a self-defined feature description procedure, which makes them invariant to a number of varied affine and imaging conditions. In this case, the results are correlated and reasoned with respect to two FR-IQA metrics (SSIM and MS-SSIM) and Pearson Coefficient observations. Since SSIM and MS-SSIM FR-IQA metrics consider the structural content of an image and the image information analyzed by MSE and NK FR-IQA metrics are somehow determined by SSIM and MS-SSIM, therefore, in the 2nd Setup only these two metrics are considered for determining the similarity index between two images. Also, the behavior similarity of the four FR-IQA metrics can be seen in Table 4.4, where all the four metrics are considered for determining the similarity index between two images (1st Setup).

Overall analysis of the experiments done are also based upon the affine invariant characteristics and computational complexity of the feature detectors.

4.6 Data Reporting

1st Setup

Dataset used for experiments: Experiments are performed on seven image sets with varied illumination and blur conditions. “Bikes”, “Tress” and “Leuven” image-sets (Image1 to Image6) are taken from the standard image dataset made available by Mikolajczyk [Mikolajczyk 2007, Appendix A.1]. The reference images (Image1 only) in image-sets labelled as ‘Ribs’ and ‘Brain-scan’ are taken from web-sources [Appendix A.1]. The distorted images in the two datasets (Ribs and Brain-scan) are created by adding Gaussian blur and by changing luminous in the reference image.

Dataset Figure.

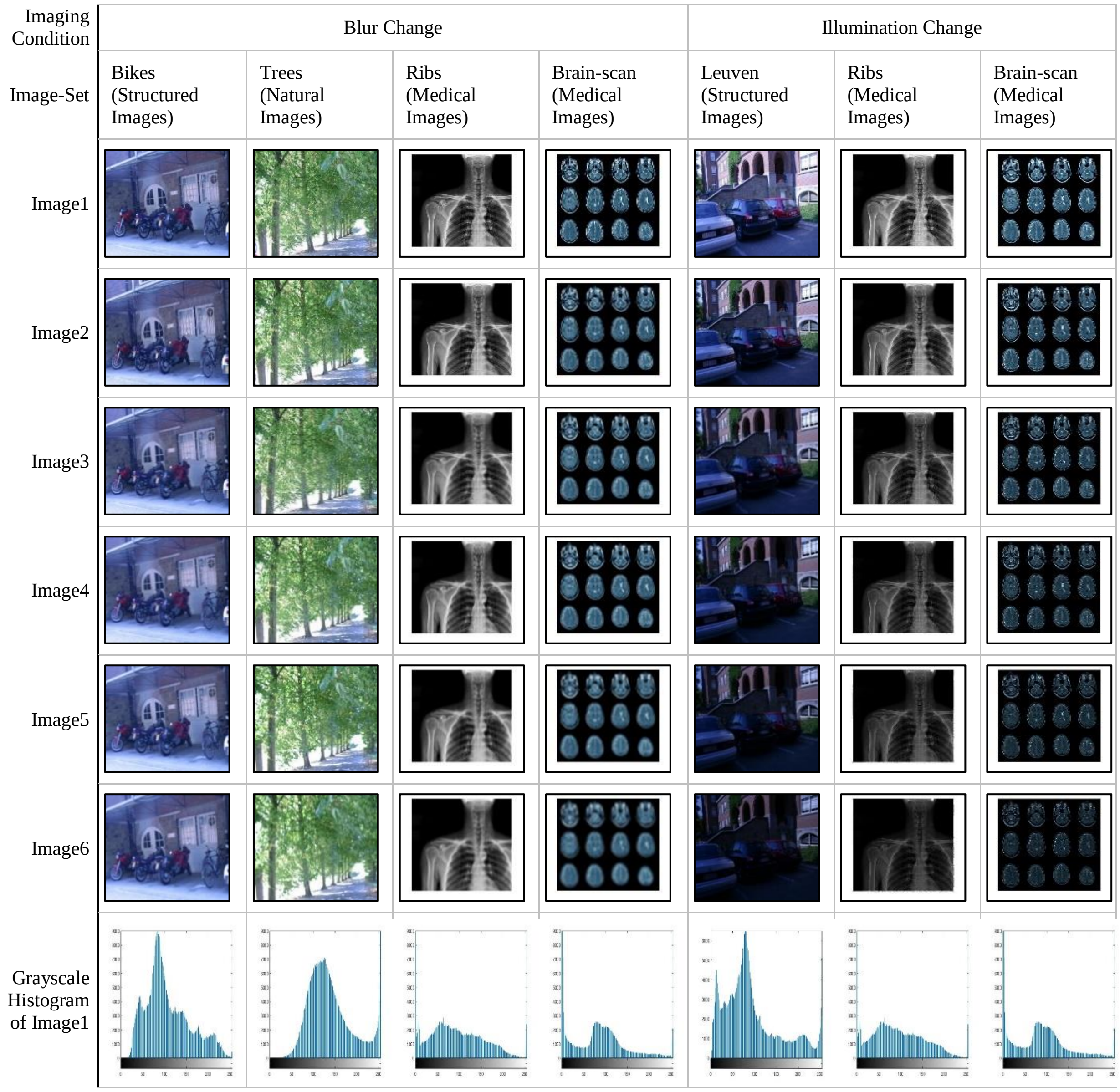


Fig. 4.1. Image-Sets with different Illumination and Blur conditions [Mikolajczyk 2007, Appendix A.1]

4.6.1 Correlation of Image Quality with performance of Image Registration Methods with respect to Keypoint Detection in an Image

From Table 4.3, it is observed that the keypoint detection for the three feature detectors can be directly correlated with the NR-IQA metric that takes into consideration the scene statistics of locally normalized luminance coefficients (BRISQUE), i.e., as the quality of image decreases, the number of keypoints detected in the image also decreases. For example: In Bikes image-set, the quality of image decreases from Image1 to Image6 (Table 4.3). In correlation, the number of keypoints detected in the image by three feature detectors (SIFT, ASIFT and SURF) also decreases from Image1 to Image6 and so is the case for other six image sets (Trees, Leuven, Ribs (Blur Change), Brain-scan (Blur Change), Ribs (Illumination Change), Brain-scan (Illumination Change)).

The keypoint detection results also relate with SSEQ metric, except for Ribs (Illumination Change), Brain-scan (Illumination Change) image-sets. Since SSEQ depends upon local spatial and spectral entropy features, it takes into account the frequency of intensity values of image pixels. So referring to grayscale histogram of Image1 (Figure 4.1) in Ribs and Brain-scan image-set, the intensity values are more towards the darker area and decreasing the illumination makes them even darker. Hence, the trend is not obvious in illumination change while considering the SSEQ metric. Moreover, the approach followed by NIQE and BLIINDS-II for quantifying the quality score of the test image did not result in the expected trend with respect to the number of keypoints detected by the three feature detectors and reasoning their behavior pattern can be considered as a potential study for future work.

Table 4.3. Keypoint Detection with respect to NR-IQA

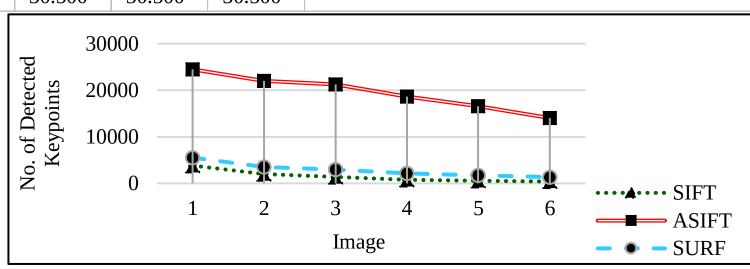
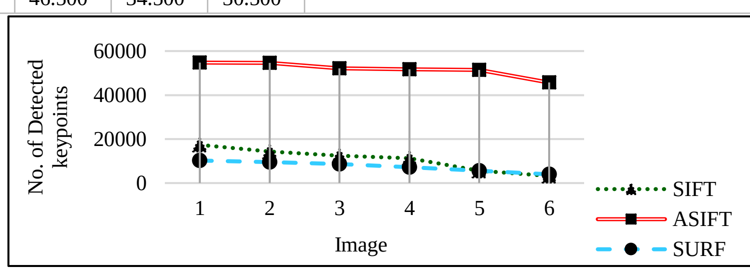
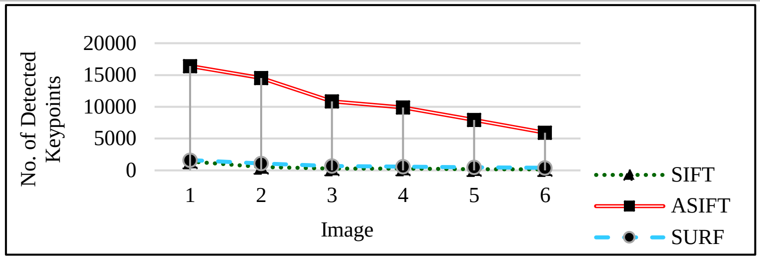
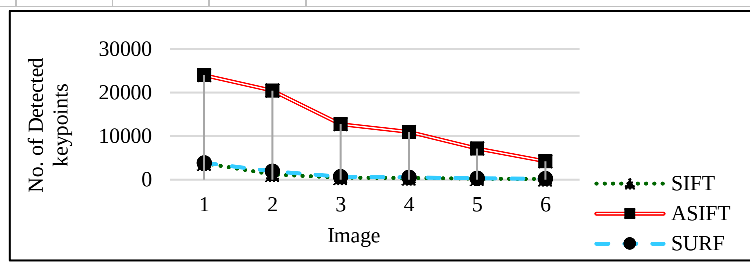
	Bikes Image Variant (Blur Change)					
	Image1	Image2	Image3	Image4	Image5	Image6
SSEQ*	24.083	40.541	44.537	48.886	50.396	52.315
NIQE*	2.627	5.093	5.851	7.842	7.998	8.251
BRISQUE*	27.278	48.172	53.719	60.184	60.533	61.496
BLIINDS-II*	12.500	34.500	35.500	30.500	30.500	30.500
						
	Trees Image Variant (Blur Change)					
	Image1	Image2	Image3	Image4	Image5	Image6
SSEQ*	11.585	17.048	24.467	40.261	46.651	48.703
NIQE*	4.741	3.8734	4.308	5.736	6.826	6.647
BRISQUE*	22.768	23.674	30.772	44.993	51.469	52.898
BLIINDS-II*	21.500	16.000	26.000	46.500	34.500	30.500
						
	Ribs Image Variant (Blur Change)					

	Image1	Image2	Image3	Image4	Image5	Image6
SSEQ*	72.132	73.364	75.516	76.511	77.523	78.547
NIQE*	3.928	5.635	7.708	8.919	9.987	10.968
BRISQUE*	40.892	65.251	69.381	79.764	86.614	90.452
BLIINDS-II*	47.000	59.500	65.500	65.500	67.500	69.500



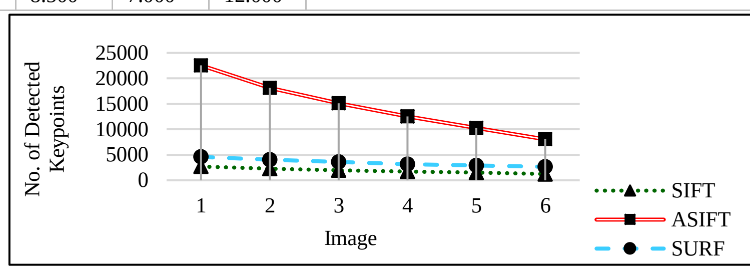
Brain-scan Image Variant (Blur Change)

	Image1	Image2	Image3	Image4	Image5	Image6
SSEQ*	59.328	64.168	72.301	72.386	78.824	78.519
NIQE*	9.896	12.010	11.892	10.301	10.918	12.161
BRISQUE*	32.601	60.871	81.472	83.503	89.583	92.537
BLIINDS-II*	54.000	61.000	63.500	63.500	67.500	65.500



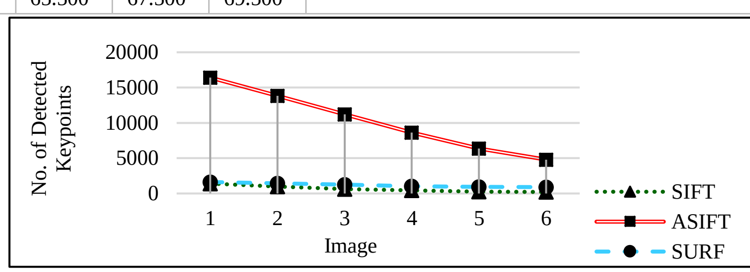
Leuven Image Variant (Illumination Change)

	Image1	Image2	Image3	Image4	Image5	Image6
SSEQ*	10.786	11.980	13.362	14.869	17.484	19.670
NIQE*	2.334	2.401	2.686	3.773	3.981	4.041
BRISQUE*	5.769	6.406	6.686	6.889	7.791	8.115
BLIINDS-II*	8.000	8.000	4.500	8.500	7.000	12.000

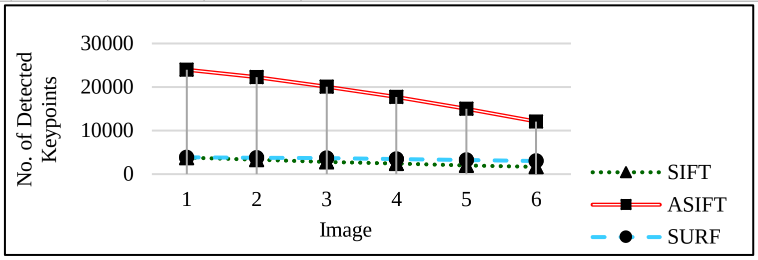


Ribs Image Variant (Illumination Change)

	Image1	Image2	Image3	Image4	Image5	Image6
SSEQ*	72.132	68.017	68.015	68.203	68.083	67.866
NIQE*	3.928	3.991	3.996	4.831	4.981	5.328
BRISQUE*	40.892	42.254	44.214	46.628	49.853	53.242
BLIINDS-II*	47.000	59.500	65.500	65.500	67.500	69.500



	Brain-scan Image Variant (Illumination Change)					
	Image1	Image2	Image3	Image4	Image5	Image6
SSEQ*	59.328	57.829	58.187	57.957	58.381	59.074
NIQE*	9.896	9.683	9.616	4.3913	4.508	4.646
BRISQUE*	32.601	33.382	34.195	34.787	36.098	37.427
BLIINDS-II*	54.000	56.500	57.500	58.000	57.000	59.000



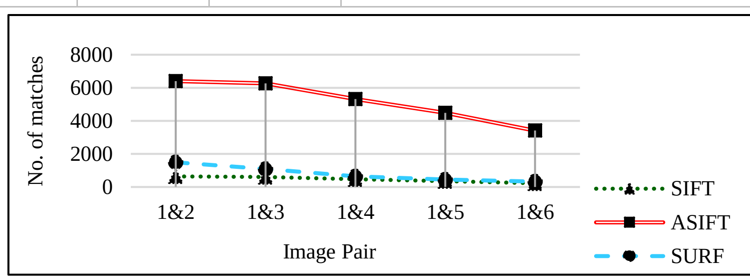
* NR-IQA metric, 0 refers to best quality and 100 refers to worst quality

4.6.2 Correlation of Image Quality with performance of Image Registration Methods with respect to Image Matching

Table 4.4 shows the result for four FR-IQA techniques (SSIM, MS-SSIM, MSE and NK). Being FR-IQA metrics, these quality metrics are used to score the image quality with respect to the reference image and hence, are used to reason how the number of matches between two images vary by the three detectors. The reference image in each image-set is chosen with respect to BRISQUE quality score of images (Table 4.3), i.e., image with lowest BRISQUE quality score is considered as the reference image. From the results tabulated in Table 4.4, it is observed that the number of correspondences between two images in all image-sets varies in accordance with the four FR-IQA metric values.

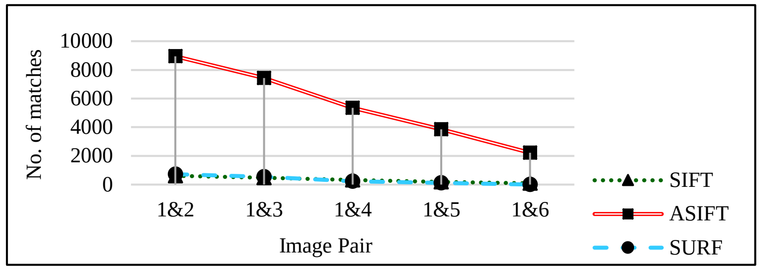
Table 4.4. Feature Matching with respect to FR-IQA

	Bikes Image Variant (Blur Change)				
	Image1&2	Image1&3	Image1&4	Image1&5	Image1&6
SSIM*	0.4727	0.4552	0.4558	0.4356	0.4114
MS-SSIM*	0.1551	0.1441	0.1401	0.1323	0.0931
MSE@	1882.7	1634.7	1856.4	1764.9	1709.7
NK#	0.9611	0.9723	0.9763	0.9787	0.9681

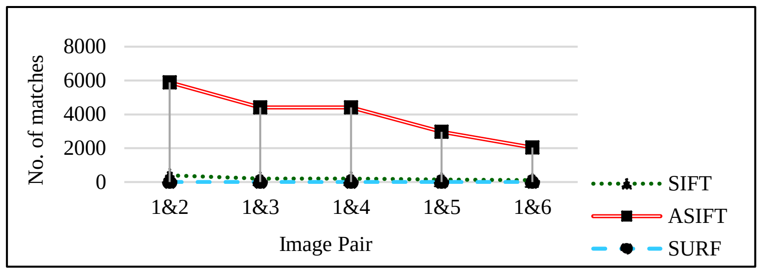


	Trees Image Variant (Blur Change)				
	Image1&2	Image1&3	Image1&4	Image1&5	Image1&6
SSIM*	0.158	0.1491	0.1329	0.1008	0.0771
MS-SSIM*	0.1058	0.0872	0.067	0.029	0.0132
MSE@	4000.4	3593.7	2868.3	2616.4	2799.2

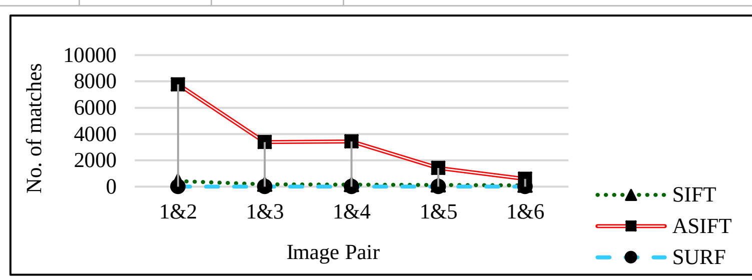
NK#	0.9183	0.9383	0.9691	0.9807	0.9883



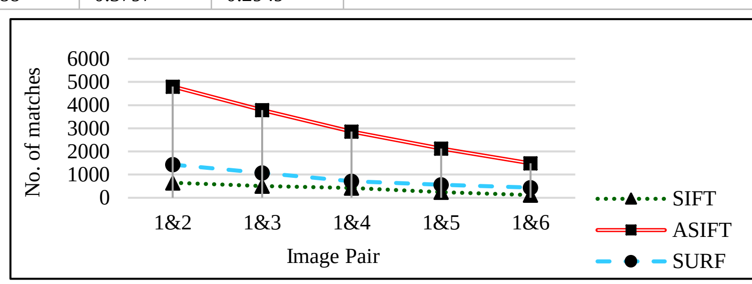
	Ribs Image Variant (Blur Change)				
	Image1&2	Image1&3	Image1&4	Image1&5	Image1&6
SSIM*	0.9561	0.9085	0.8801	0.8731	0.8503
MS-SSIM*	0.9911	0.9708	0.9688	0.9450	0.9189
MSE@	10.693	30.501	30.501	53.463	76.134
NK#	0.9990	0.9976	0.9976	0.9961	0.9947



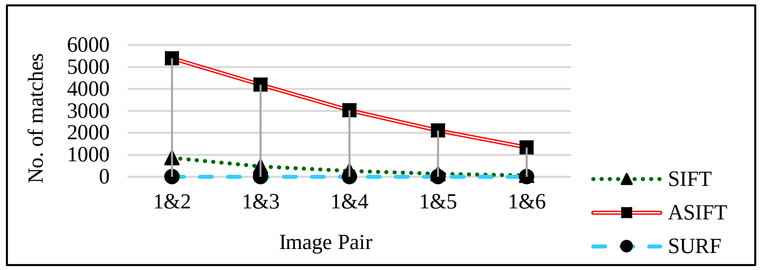
	Brain-scan Image Variant (Blur Change)				
	Image1&2	Image1&3	Image1&4	Image1&5	Image1&6
SSIM*	0.8623	0.7920	0.7419	0.6801	0.6360
MS-SSIM*	0.9615	0.8908	0.8769	0.8033	0.7449
MSE@	218.3	471.6	471.6	639.1	749.8
NK#	0.9842	0.9720	0.9720	0.9643	0.9589



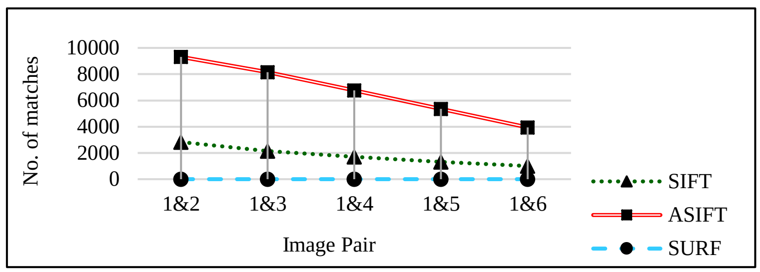
	Leuven Image Variant (Illumination Change)				
	Image1&2	Image1&3	Image1&4	Image1&5	Image1&6
SSIM*	0.3895	0.2931	0.2456	0.2121	0.1742
MS-SSIM*	0.5589	0.347	0.1905	0.1812	0.0849
MSE@	2121.4	3490.4	4977.2	5753.2	7162.6
NK#	0.6953	0.563	0.4488	0.3797	0.2949



	Ribs Image Variant (Illumination Change)				
	Image1&2	Image1&3	Image1&4	Image1&5	Image1&6
SSIM*	0.9859	0.9481	0.8927	0.8311	0.7710
MS-SSIM*	0.9711	0.8913	0.7800	0.6604	0.5498
MSE@	160.6	562.5	1097.9	1690.2	2291.4
NK#	0.9596	0.9237	0.8924	0.8654	0.8424



	Brain-scan Image Variant (Illumination Change)				
	Image1&2	Image1&3	Image1&4	Image1&5	Image1&6
SSIM*	0.9885	0.9576	0.9120	0.8610	0.8120
MS-SSIM*	0.9696	0.8835	0.7607	0.6265	0.5014
MSE@	95.4	335.5	661.2	1028.8	1407.1
NK#	0.9724	0.9475	0.9252	0.9056	0.8887



*Value range: [-1,1] 1= identical set, @ Value range: [-1,1] 1= perfect match, -1= completely anti-correlated, 0=completely uncorrelated, #Value range: [0,∞) Low MSE=Low error

2nd Setup

Dataset used for experiments: Experiments are done on the standard image dataset made available by Mikolajczyk [Mikolajczyk 2007, Appendix A.1]. Figure 4.2 shows all the images in the dataset as per the classification of image-sets given by Mikolajczyk.

Different imaging conditions contained in the dataset: The complete dataset consists of total 48 images which are grouped in eight image-sets each containing six images. Each image-set defines alterations of an image scene in five different imaging condition: 1) Viewpoint change (Image-set: Graffiti and Wall), 2) Scale change (Image-set: Boat and Bark), 3) Image blur (Image-set: Bikes and Trees), 4) Illumination change (Image-set: Leuven) and 5) JPEG compression (Image-set: Ubc). Viewpoint change, scale change and image blur conditions have been applied to two image sets each. One image set contains structured images and the other comprises of natural images. For example: two image-sets for viewpoint change are Graffiti (containing structure images) and Wall (containing natural images).

Dataset Figure.

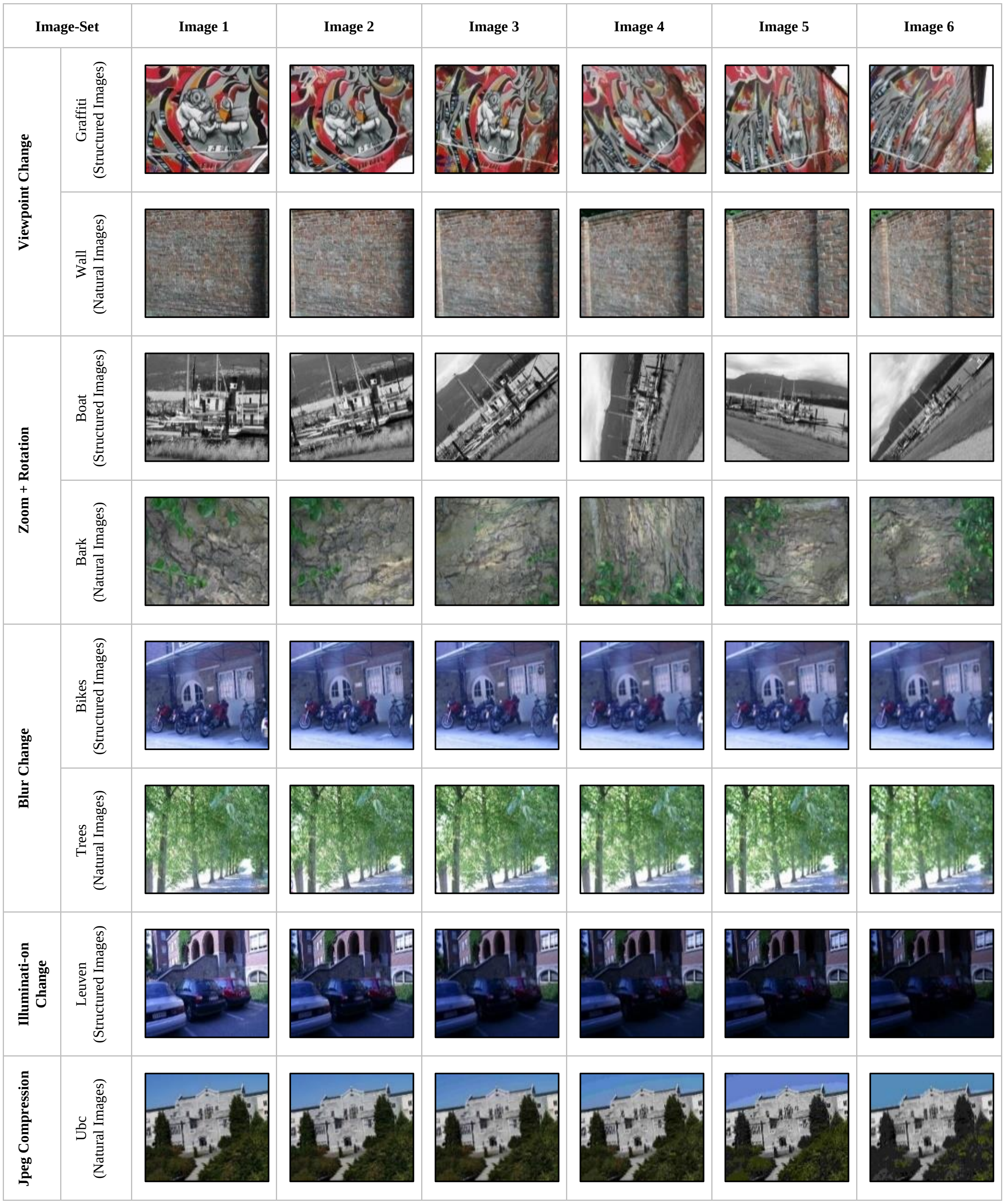


Fig. 4.2. Mikolajczyk Dataset of 48 images containing eight Image-Sets (each containing six images) under different imaging conditions [Mikolajczyk 2007, Appendix A.1]

Pearson Coefficient: To understand the correlation between every pair of image in each image-set, Pearson Coefficient (correlation coefficient) is used. The results are tabulated in Table 4.5 and are graphically represented in Figure 4.3 for clear interpretation.

Table 4.5. Pearson Coefficient* for each pair of image in each Image-Set

Image Pair	Image-Set							
	Graffiti	**Wall**	**Boat**	**Bark**	**Bikes**	**Trees**	**Leuven**	**Ubc**
1&2	0.0532	0.1068	0.0920	0.1607	0.5267	0.1991	0.6261	0.9668
1&3	0.0134	0.1186	-0.1002	-0.0854	0.5724	0.2358	0.5077	0.9472
1&4	0.0110	0.0880	-0.0633	-0.0625	0.5173	0.3189	0.3417	0.9117
1&5	-0.0339	0.0794	-0.1835	0.0913	0.5276	0.3527	0.2912	0.8686
1&6	-0.0037	0.0831	-0.0634	-0.0815	0.5222	0.3314	0.2104	0.8459
2&3	-0.0061	0.2137	-0.0502	-0.0511	0.6973	0.3565	0.7080	0.9593
2&4	0.1038	0.1606	-0.1168	-0.0497	0.6550	0.2767	0.4653	0.9235
2&5	-0.0051	0.1535	-0.1317	0.1912	0.6905	0.2475	0.3561	0.8801
2&6	-0.0271	0.1638	-0.1464	-0.0930	0.6735	0.2741	0.2566	0.8573
3&4	0.0383	0.2165	0.0192	0.1951	0.7995	0.3480	0.5497	0.9417
3&5	0.0404	0.1837	0.1463	0.0814	0.8221	0.2983	0.4366	0.8979
3&6	0.0011	0.1986	-0.0391	0.1408	0.8039	0.3448	0.2954	0.8749
4&5	0.0257	0.2354	0.0168	0.0590	0.8716	0.4160	0.4884	0.9253
4&6	-0.0238	0.2188	0.2591	0.0772	0.9194	0.5319	0.3250	0.9035
5&6	0.0070	0.2545	0.0031	-0.0030	0.9205	0.5545	0.4830	0.9353

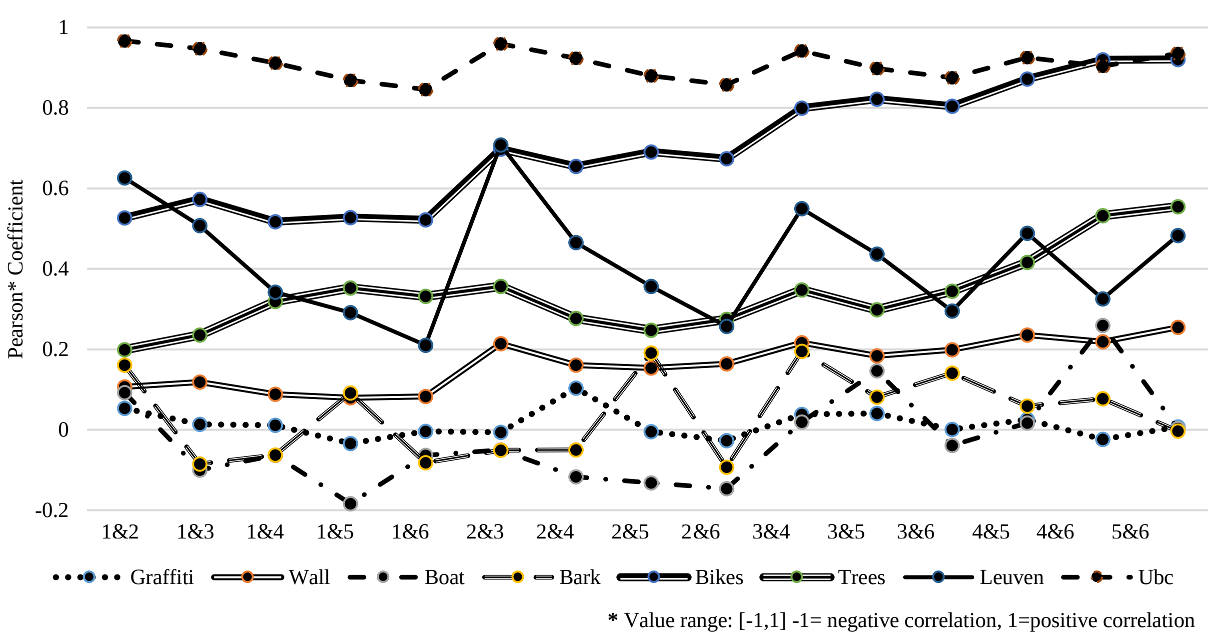


Fig. 4.3. Graphical Representation of Pearson Coefficient for every pair of image in each Image-Set

4.6.3 Correlation of Image Quality with performance of Image Registration Methods with respect to Keypoint Detection in an Image

Figure 4.4 and Figure 4.5 displays the images with detected keypoints by the six feature detectors for two images: Image1 and Image6, for two image sets: Graffiti and Leuven. The keypoints detected in these figures are represented using an ellipse, outlined using green color. The keypoint detection output for six feature detectors is analyzed with respect to the output of SSEQ and BRISQUE NR-IQA metric values of respective image. The results are collectively tabulated in Table 4.6. Reasoning for quality assessment metric value and keypoint detection procedure is summarized in Table 4.7.

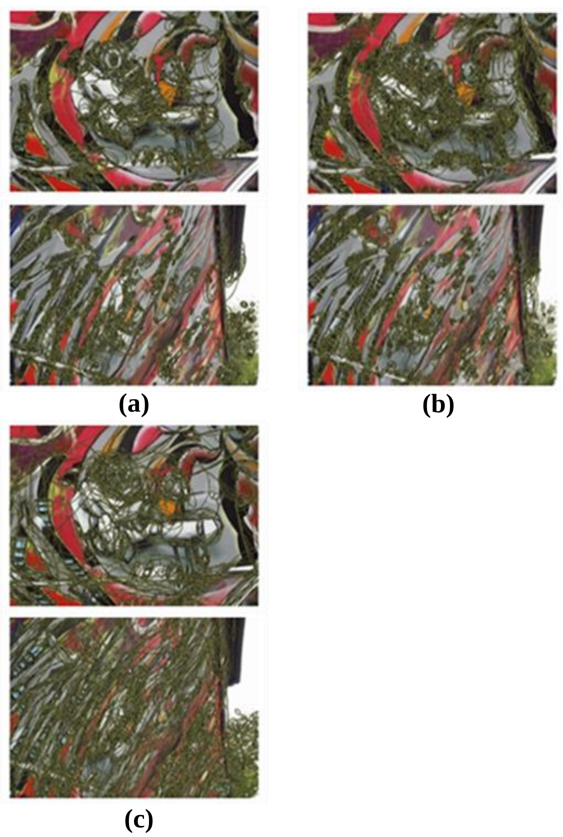


Fig. 4.4. Keypoint detection done for two images in Graffiti image-set (Image1 and Image6) using (a) Harris-Affine (shown) (b) Hessian Affine (shown) (c) MSER (shown) (d) SIFT (e) ASIFT (f) SURF. The detectors find respectively 1758, 2454, 533, 3094, 28435, 3961 number of interest points for Image1 and 1846, 1845, 896, 5162, 35039, 5591 number of interest points for Image6

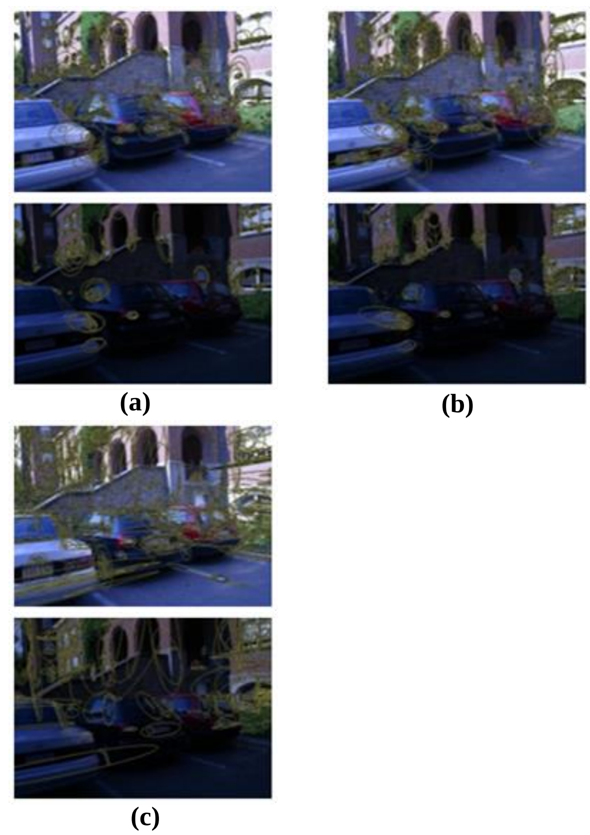


Fig. 4.5. Keypoint detection done for two images in Leuven image-set (Image1 and Image6) using (a) Harris-Affine (shown) (b) Hessian Affine (shown) (c) MSER (shown) (d) SIFT (e) ASIFT (f) SURF. The detectors find respectively 902, 1125, 566, 2709, 22556, 4590 number of interest points for Image1 and 329, 439, 249, 1240, 8044, 2653 number of interest points for Image6

Table 4.6. Keypoint Detection with respect to NR-IQA

Image-Set	Image	SSEQ*	BRISQUE*	Number of Detected Keypoints					
				Harris affine	Hessian Affine	MSER	SIFT	ASIFT	SURF
Graffiti	Image1	13.822	4.966	1758	2454	533	3094	28435	3961
	Image2	16.749	8.630	1973	2731	596	3539	31939	4088
	Image3	18.626	12.182	2172	2784	659	3982	35126	4539
	Image4	14.847	6.197	1976	2296	682	4116	33494	4706
	Image5	23.642	18.950	2153	2434	786	4493	35679	5066
	Image6	21.209	14.899	1846	1845	896	5162	35039	5591
Wall	Image1	20.969	23.376	2267	1375	2363	10612	35020	11697
	Image2	30.720	26.851	2013	1376	1921	11860	39237	9040
	Image3	30.081	26.356	1969	1343	1905	11746	39204	8778
	Image4	26.831	27.116	2088	1449	1873	11363	41379	8683
	Image5	27.914	25.822	2165	1537	1902	11864	42186	8669
	Image6	30.461	25.333	2228	1524	1884	11632	41125	8655
Boat	Image1	12.769	21.980	3023	3146	1524	9688	46226	6208
	Image2	16.788	18.327	2935	2972	1456	9278	45089	6621
	Image3	19.240	3.041	2379	2587	1151	7114	38695	5578
	Image4	8.162	15.492	1423	1433	725	5670	25970	4974
	Image5	21.390	18.831	1199	1217	653	5376	24861	4242
	Image6	6.058	1.351	1018	1066	562	4510	22777	4528
Bark	Image1	33.824	33.163	421	493	431	4226	36504	3719
	Image2	33.703	34.015	277	352	266	3465	36746	3831
	Image3	26.939	23.974	294	339	407	4589	33672	4247
	Image4	31.003	26.661	519	510	704	5231	38287	4602
	Image5	22.653	22.827	520	510	684	4811	35530	4501
	Image6	19.510	23.941	586	495	780	5008	36988	4906
Bikes	Image1	24.083	27.278	878	1025	606	3825	24488	5546
	Image2	40.541	48.172	665	926	350	2015	21992	3528
	Image3	44.537	53.719	624	920	293	1397	21265	2943
	Image4	48.886	60.184	482	765	196	802	18626	2145
	Image5	50.396	60.533	384	655	153	571	16547	1731
	Image6	52.315	61.496	304	513	106	407	14034	1351
Trees	Image1	11.585	22.768	5504	3685	2663	14290	52197	9600
	Image2	17.048	23.674	5516	3782	2832	12381	51397	10233
	Image3	24.467	30.772	5617	4037	2522	17354	54586	8640
	Image4	40.261	44.993	4684	4217	2109	11343	54753	7155
	Image5	46.651	51.469	3337	4094	1708	5648	51669	5661
	Image6	48.703	52.898	2064	3414	1052	3238	45675	3978

Image-Set	Image	SSEQ*	BRISQUE*	Number of Detected Keypoints					
				Harris affine	Hessian Affine	MSER	SIFT	ASIFT	SURF
Leuven	Image1	10.786	5.769	902	1125	566	2709	22556	4590
	Image2	11.980	6.406	723	953	480	2294	18156	4050
	Image3	13.362	6.686	615	805	416	1969	15132	3599
	Image4	14.869	6.889	500	664	346	1719	12551	3167
	Image5	17.484	7.791	399	558	292	1532	10298	2899
	Image6	19.670	8.115	329	439	249	1240	8044	2653
Ubc	Image1	10.676	9.788	1402	1570	770	5762	33633	6405
	Image2	35.336	30.647	1425	1565	794	6989	33385	6134
	Image3	44.202	46.237	1421	1579	809	8174	33754	6082
	Image4	56.359	63.490	1400	1615	874	7796	34397	5849
	Image5	58.673	87.554	1540	1647	1166	5563	36273	5019
	Image6	58.789	87.187	1368	1728	929	3568	36549	3767

*0 refers to best quality and 100 refers to worst quality

Table 4.7. Result Analysis from Table 4.6.

Imaging Variations		Results after analyzing Table 4.6	Reasoning
Viewpoint Change (Viewpoint angle increases from Image1 to Image6)	Graffiti	**Quality:** Not in any persistent order **Harris Affine, Hessian Affine and ASIFT:** Number of keypoints detected oppositely coincides with how the quality of image varies i.e. as the image quality decreases, number of keypoints increases. **MSER, SIFT and SURF:** Keypoints Increases as we go from Image1 to Image6.	**Quality:** Not in any persistent order as the images contain objects captured at different viewpoint, so quality can vary due to various features. **Detected Keypoints:** No consistent order as there is inconsistent correlation between images.
	Wall	**Quality:** Not in any persistent order Harris Affine, Hessian Affine and SURF : Keypoints Decreases as we go from Image1 to Image6 irrespective of Image Quality **MSER, SIFT and ASIFT:** There is no consistent order in the detected keypoints neither individually nor with respect to the NR-IQA method.	
Zoom + Rotation	Boat	**Quality:** Not in any persistent order **Harris Affine, Hessian Affine, MSER, SIFT and ASIFT**: Keypoints Increases as we go from Image1 to Image6 **SURF:** There is no consistent order either in the detected keypoints individually nor with respect to the NR-IQA method.	**Quality:** Not in any persistent order as the images contain objects captured at different viewpoint, so quality can vary due to various features. **Detected Keypoints:** No consistent order as there is

Imaging Variations		Results after analyzing Table 4.6	Reasoning
	Bark	**Quality:** Increases as we go from Image1 to Image6 **All detectors work in the same manner** There is no consistent order either in the detected keypoints individually nor with respect to the NR-IQA method.	inconsistent correlation between images.
Blur (Blur factor increases from Image1 to Image6)	Bikes	**Quality:** Decreases as we go from Image1 to Image6 **All detectors work in the same manner**. As Image Quality decreases, number of keypoints detected also decreases.	**Quality:** Decreases as we go from Image1 to Image6 as the blur effect increases in the same manner. **Detected Keypoints:** For Bikes: There is a consistent order where the detected keypoints decreases as the image quality decreases because of the persistent correlation between images (check for pairs 1&2, 2&3, 3&4, 4&5, 5&6). For Trees: No consistent order as there is inconsistent correlation between images.
	Trees	**Quality:** Decreases as we go from Image1 to Image6 **All detectors work in the same manner**. There is no consistent order either in the detected keypoints individually nor with respect to the NR-IQA method.	
Illumination Change (Illumination decreases from Image1 to Image6)	Leuven	**Quality:** Decreases as we go from Image1 to Image6 **All detectors work in the same manner.** As Image Quality decreases, number of keypoints detected also decreases.	**Quality:** Decreases as we go from Image1 to Image6 as the illumination decreases in the same manner. **Detected Keypoints:** There is a consistent order where the detected keypoints decreases as the image quality decreases because of the persistent correlation between images.
Jpeg Compression (Compression ratio increases as we go from Image1 to Image6.)	Ubc	**Quality:** Decreases as we go from Image1 to Image6 **Harris Affine, Hessian Affine, MSER, SIFT and ASIFT:** There is no consistent order in the detected keypoints neither individually nor with respect to the NR-IQA method. **SURF:** As Image Quality decreases, number of keypoints detected also decreases.	**Quality:** Decreases as we go from Image1 to Image6 as the compression ratio increases in the same manner. **Detected Keypoints:** No consistent order as there is inconsistent correlation between images (check for pairs 1&2, 2&3, 3&4, 4&5, 5&6).

4.6.4 Correlation of Image Quality with performance of Image Registration Methods with respect to Image Matching

Table 4.8 shows the result for SSIM and MS-SSIM. These image quality metrics are used to identify how the number of matches between two images vary as the similarity value varies between them. The reference image in each image set is chosen with respect to the NR-IQA value of image in Table 4.6. For example: In case of Boat and Bark image set, Image6 has the lowest quality score value (for both SSEQ and BRISQUE metric, indicating the best quality of image in the image-set). So, when matching is done in these image sets, Image6 is always taken as the reference image.

For this comparative evaluation, SIFT, ASIFT and SURF feature detector is selected as their descriptors are invariant to a number of imaging conditions and there exists substantial references in literature that proves their efficiency over other feature descriptors [Yu and Morel 2011]. The results show that for all image-sets, number of matches between two images decreases as the similarity index decreases and vice-versa except for two image-sets contained in Zoom + Rotation imaging condition where neither the FR-IQA values nor the number of matches detected by detectors are consistent. The reason behind this is the discrepancy between zoom and rotation factors in adjacent images, resulting in inconsistency of results. Table 4.9 confines the correlation coefficient between SSIM and MS-SSIM values for each image-set and shows least correlation for Boat and Bark image-set, hence reasoning the inconsistency in number of correspondences between images. Figure 4.6 and 4.7 shows the image matching results for Graffiti and Leuven image-set by the three detectors (straight lines between the images represent correspondences).

Table 4.8. Feature Matching with respect to FR-IQA

Image-Set	Image Pair	SSIM	MS-SSIM	Number of Matches & Time Taken (in Sec)					
				SIFT		ASIFT		SURF	
				# Matches	Time	# Matches	Time	# Matches	Time
Graffiti	Image1&2	0.1983	0.014	315	3.88	3365	131.56	236	3.09
	Image1&3	0.1836	0.0041	76	3.52	2299	131.26	59	3.03
	Image1&4	0.1822	0.0012	40	3.54	1509	131.23	15	3.00
	Image1&5	0.1601	0.0009	2	3.75	854	130.01	8	2.89
	Image1&6	0.1578	0.0008	0	3.28	447	130.02	5	3.02
Wall	Image1&2	0.116	0.0327	1169	5.68	11142	139.28	1704	4.89
	Image1&3	0.1177	0.0458	691	5.99	7475	139.65	891	4.99
	Image1&4	0.0977	0.0136	209	5.86	3981	140.34	258	4.83
	Image1&5	0.1018	0.0163	13	5.99	2294	139.78	29	4.78

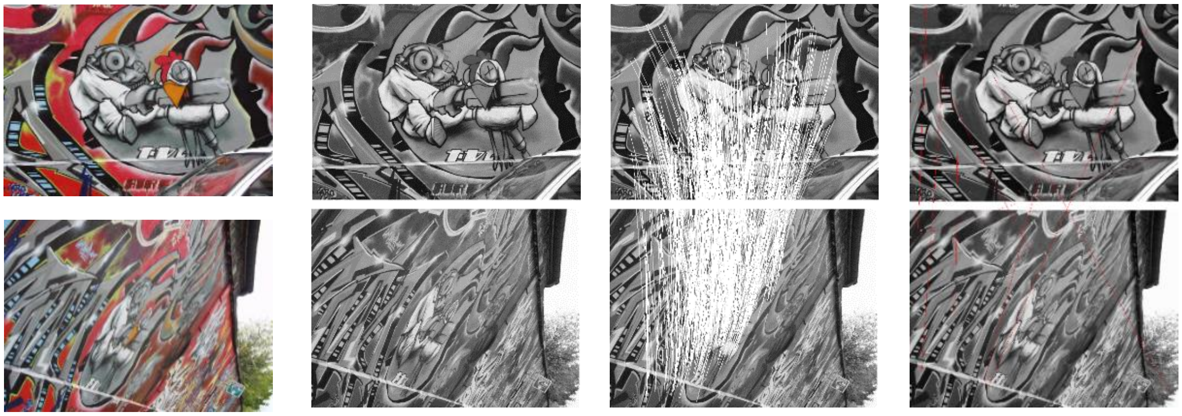
Image-Set	Image Pair	SSIM	MS-SSIM	Number of Matches & Time Taken (in Sec)					
				SIFT		ASIFT		SURF	
				# Matches	Time	# Matches	Time	# Matches	Time
	Image1&6	0.1023	0.0216	0	5.82	801	139.12	1	4.88
Boat	Image6&1	0.173	0.0029	0	5.25	33	139.26	4	4.25
	Image6&2	0.1839	0.0103	0	5.45	89	138.26	8	4.12
	Image6&3	0.2513	0.0147	47	5.33	351	138.28	19	4.29
	Image6&4	0.3121	0.0563	0	5.41	388	139.33	16	4.36
	Image6&5	0.2974	0.0099	45	5.55	994	138.18	18	4.29
Bark	Image6&1	0.1889	0.0029	12	3.49	97	129.36	5	3.25
	Image6&2	0.1960	0.0103	15	3.32	199	130.56	33	3.44
	Image6&3	0.1962	0.0147	212	3.49	1456	129.52	69	3.52
	Image6&4	0.1634	0.0563	91	3.39	1389	130.99	164	3.32
	Image6&5	0.1644	0.0099	228	3.55	2025	130.55	287	3.49
Bikes	Image1&2	0.4727	0.1551	640	2.56	6402	98.65	1492	1.89
	Image1&3	0.4552	0.1441	603	2.69	6263	98.35	1092	1.99
	Image1&4	0.4558	0.1401	469	2.98	5315	98.52	649	1.82
	Image1&5	0.4356	0.1323	363	2.78	4488	99.33	459	1.75
	Image1&6	0.4114	0.0931	263	2.89	3408	98.26	326	1.65
Trees	Image1&2	0.158	0.1058	614	6.28	8946	142.29	723	4.12
	Image1&3	0.1491	0.0872	473	6.23	7439	142.53	542	4.28
	Image1&4	0.1329	0.067	310	6.53	5359	142.39	240	4.96
	Image1&5	0.1008	0.029	186	6.21	3845	142.33	123	4.28
	Image1&6	0.0771	0.0132	91	6.07	2220	142.03	5	4.36
Leuven	Image1&2	0.3895	0.5589	649	1.67	4803	56.23	1431	1.06
	Image1&3	0.2931	0.347	509	1.58	3787	51.26	1067	1.25
	Image1&4	0.2456	0.1905	428	1.36	2857	47.18	715	0.98
	Image1&5	0.2121	0.1812	241	1.41	2121	45.29	565	0.92
	Image1&6	0.1742	0.0849	118	1.05	1488	44.12	440	0.96
Ubc	Image1&2	0.9023	0.987	1224	4.99	10492	132.15	3428	3.89
	Image1&3	0.8443	0.9715	1127	4.86	10366	132.33	2255	3.88
	Image1&4	0.7543	0.9345	885	4.69	9250	132.45	1021	3.79
	Image1&5	0.6291	0.8435	485	4.88	6779	132.26	776	3.92
	Image1&6	0.5148	0.7352	246	4.97	4397	132.45	309	3.99

* Value range: [-1,1] 1= identical set

Table 4.9. Pearson Coefficient* between SSIM and MS-SSIM values

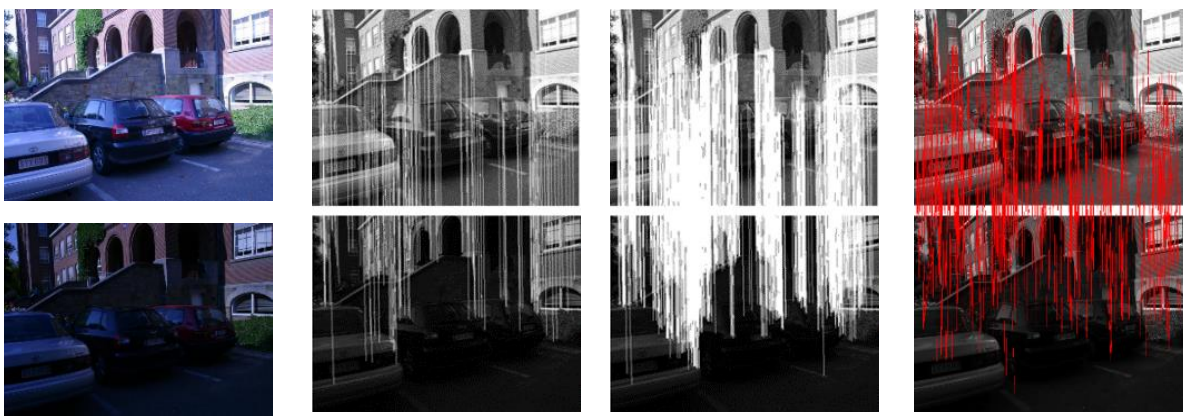
Graffiti	Wall	Boat	Bark	Bikes	Trees	Leuven	Ubc
0.809438304	0.9498416	0.67801919	-0.58366059	0.96548553	0.98982625	0.9907789	0.98212497

* Value range: [-1,1] -1= negative correlation, 1=positive correlation



(a) **(b)** **(c)** **(d)**

Fig. 4.6. With respect to Original Pair of Image (a), robustness to Viewpoint Change by: (b) SIFT (c) ASIFT (d) SURF find respectively 0, 447 And 5 correct matches between Image1 and Image6 in Graffiti Image-Set



(a) **(b)** **(c)** **(d)**

Fig. 4.7. With respect to Original Pair of Image (a), robustness to Illumination Change by: (b) SIFT (c) ASIFT (d) SURF find respectively 118, 1488 And 440 correct matches between Image1 and Image6 in Leuven Image-Set

4.7 Result Analysis and Interpretation

Feature detectors could be analyzed with respect to their efficiency in terms of stable number of keypoints detected in an image, correct number of correspondences found between an image pair under extreme changing imaging conditions and its computational complexity. Keeping all these criteria in mind, few widely used feature detectors are compared and analyzed in this chapter under different imaging conditions and while dealing with varied quality of images, arguing their applicability in AR applications.

The comparative analysis is carried out in two setups, where in 1st Setup, only blur and illumination imaging conditions of medical, natural and structured images are taken into consideration for comparative analysis of three widely used feature detectors i.e. SIFT, ASIFT and SURF. Performance of these detectors is also analyzed with respect to the quality of images used for determining the two imaging conditions. IQA metrics are utilized in two forms: 1) NR_IQA metrics (SSEQ, NIQE, BRISQUE and BLIINDS-II) are used for interpreting the keypoint detection behavior of SIFT, ASIFT and SURF in images (Table 4.3). 2) FR-IQA metrics (SSIM, MS-SSIM, MSE and NK) are used for interpreting the feature matching performance of the three detectors (Table 4.4). Also, NR-IQA metrics quality scores of images helps in selecting the reference image in each image-set for performing FR_IQA and image matching (Table 4.4). Results show that the number of keypoints detected in an image and the number of matches found between two images can be associated with the quality of that image and the similarity index value between two images respectively in both blur and illumination imaging condition.

In the 2nd Setup, relatable outcome of 1st Setup is treated as the motivation for extending the comparative analysis for six feature detectors (Harris-Affine, Hessain-Affine, MSER, SIFT, ASIFT and SURF) under five varying imaging conditions: viewpoint change, scale change, blur change, illumination change and JPEG compression. As done in 1st Setup, the quality of images in this setup too is determined using IQA metrics in two forms: 1) NR-IQA metrics are used for interpreting the keypoint detection behavior of all the six feature detectors (Table 4.6). The NR-IQA metrics used in this case, however, are SSEQ and BRISQUE as the quality scores prediction of images by NIQE and BLIIND-II NR-IQA metrics in 1st Setup were not determinable (Section 4.6.1 discusses the behavior analysis of SSEQ, NIQE, BRISQUE and BLIINDS-II NR-IQA metrics). 2) FR-IQA metrics are used for interpreting the feature matching performance of three feature detectors, SIFT, ASIFT and SURF (Table 4.8). Only these three feature detectors are used for feature matching analysis because they have a self-defined feature description procedure of their own, which makes it possible to perform feature matching among the extracted features. Also, FR-IQA metrics used in this setup are SSIM and MS-SSIM as the image information used by MSE and NK FR-IQA metrics (used in 1st Setup) are in some form incorporated by the two chosen metrics. In addition, the behavior similarity of four FR-IQA metrics (SSIM, MS-SSIM, MSE and NK) can be seen in Table 4.4. The 2nd Setup also involves Pearson Coefficient to study the correlation between every pair of image in each image-set and the two FR-IQA metrics results to determine the traits of the six feature detectors in terms of number of detected keypoints in an image and number of matches between two images with respect to image quality (Table 4.5, Table 4.9).

Experimental results show that for keypoint detection in an image, performance of six feature detectors, in most cases, can be correlated with the quality of images used for experimentation (Table 4.7 gives the detailed reasoning for performance of the six feature detectors in terms of keypoint detection in an image for all image-sets). It can be always said that the number of keypoints detected depends on the sharpness and clarity of the image, and so, as depicted in Table 4.6, for Bikes and Leuven image-set that represents blur and illumination change (sharpness and clarity) in a structured image scene respectively, image quality decreases as the illumination decreases and blur increases in subsequent images and accordingly the number of keypoints detected in images also varies with noticeable correlation. Number of detected keypoints in an image cannot be related to image degradation like JPEG compression, however, for Ubc image-set where JPEG compression ratio increases from Image1 to Image6 and thus quality decreases in the same order, SURF detector follows the same trend. For feature matching, it is seen that, for all image-sets except Boat and Bark image-sets, the number of correspondences between two images decreases as the similarity index decreases between them and vice-versa. The unrelatable performance of the feature detectors for Boat and Bark image-sets is reasoned with respect to Pearson Coefficient (Section 4.6.4).

Interpreting the applicability of the six feature detectors in an AR system, it can be seen that even though ASIFT is able to detect a noticeable higher amount of keypoints in an image and also outperforms both SIFT and SURF in terms of feature matching efficiency, the computational complexity of the detector is much more than what could be accepted for AR applications. Also, SURF performs better than SIFT in accuracy and has a lower computational complexity than ASIFT, but its processing speed is not appreciable for a real time AR system. Therefore, registration process in AR applications could be made more accurate and computationally less expensive by optimizing the combination of an appropriate keypoint detection algorithm with a suitable descriptor. As an example, MSER detector in all cases behaves as SIFT and ASIFT detector for different imaging conditions (Table 4.7). So to make the detector more affine invariant, it can be combined with SIFT or SURF descriptor for better results and could be used for further formulations of pose recovery (motivation for the proposed improvement in image registration method for AR discussed in Chapter 6).

4.8 Summary

In this chapter, comparative analysis of some widely used feature detectors with respect to different imaging conditions in terms of viewpoint change, scale change, image blur, illumination change and JPEG compression is discussed. Reasoning of the performance behavior of these detectors is weighed upon the quality of images they are being processed upon and Pearson coefficient. Next chapter presents an improved NR-IQA Model and a No-Reference Video Quality Assessment Model Based on frame analysis to analyze different features that deteriorate the quality of an image/video and to predict the distortion present in an image in a more efficient manner.