Iraqi Journal of Science
ISSN: 0067-2904
GIF: 0.851

New multispectral images classification method based on MSR and Skewness implementing on various sensor scenes

Taghreed A. H. Naji *
Department of Physics, College of Education for Pure Sciences Ibn Al-Haitham,University of Baghdad, Baghdad, Iraq

Abstract
A new features extraction approach is presented based on mathematical form the modify soil ratio (MSR) and skewness for numerous environmental studies. This approach is involved the investigate on the separation of features using frequency band combination by ratio to estimate the quantity of these features, and it is exhibited a particular aspect to determine the shape of features according to the position of brightness values in a digital scenes, especially when the utilizing the skewness. In this research, the marginal probability density function G(MSR) derivation for the MSR index is corrected, that mentioned in several sources including the source (Aim et al.). This index can be used on original input features space for three different scenes, and then implemented the marginal probability density function of MSR values to stretch the histograms of MSR images without any processing. Skewness is proposed on MSR images and combined with multispectral bands of original scene for land cover classification. This is a new method for extensively observing the types of features and its changes. The Hyperion data were utilized in this work; because they contain abundant details information for distinguish the different types of features.

Keywords: MSR index, probability theory, skewness statistical, Hyperion hyperspectral imager, unsupervised classification.

طريقة جديدة لتصنيف الصور متعددة الاطياف استنادا إلى تطبيق (MSR) والالتواء لمشاهدِ المتحسّسات المُخْتَلِفةِ

تغريد عبد الحميد ناجي *
قسم الفيزياء، كلية التربية للعلوم الصرفة ابن الهيثم، جامعة بغداد، بغداد، العراق

الخلاصة
لقد تم تقديم نهجا ً جديدا ً لاستخراج المعالم استنادا إلى الشكل الرياضي لتعديل نسبة التربة(MSR) والإلتواء للعديد من الدراسات البيئية. استلزم هذا النهج على تحقيق الفصل بين المعالم باستعمال تردد ترابط الحزم عن طريق حساب النسبة بين الحزم لتَخمين كميةِ المعالم، وقد أظهرت جانب معين لتحديد شكل المعالم وفقا لقيم السطوع في المشاهد الرقمية، وخصوصا عند إستعمال إحصاء الإلتواء. في هذا البحث تم تصحيح اشتقاق معادلة دالة كثافة الاحتمالية الحّدية G(MSR) لمؤشر تعديل نسبة التربة (MSR) الذي ذَكرَ في عِدّة مصادر، من ضمنها المصدرِ (Aim et al.). يمكن إستعمال المؤشر (MSR) على مدخلات فضاء المعالم لثلاث مشاهد أصلية مختلفة، ومن ثم طُبقت دالة كثافة الاحتمالية الحّدية لقيم (MSR) لتوسيع المخطط

*Email: taghreedaltaie@yahoo.com

التكراري لصور (MSR) بدون أيّ معالجة. اُقترح إحصاء الإلتواء على الصور (MSR) ومن ثم اُدمجت النتيجة مع الحزم متعددة الأطياف للمشهد الاصلي لتصنيف الغطاء الأرضي. هذه طريقة جديدة لمراقبة أنواع المعالم وتغيراتها على نطاق واسع. تم استعمال بيانات الهايبرون في هذا العمل، لأنها تحتوي على تفاصيل معلومات وفيرة للتمييز بين أنواع مختلفة من المعالم.

الكلمات المفتاحية: مؤشر (MSR)، احصاء الالتواء، نظرية الاحتمالية، تصوير فائق الأطياف الهايبرون، التصنيف الغير موجه.

Introduction

Driven by the need in ecological, climate and many other studies to quantitatively assess plants conditions from remote sensing measurements, numerous vegetation indices have been developed using the measurements in red (or visible) and near infrared (NIR) bands. In environmental researches, band ratios are frequently used in order to quantify the amount of plants that may appear in a multispectral digital image. Such an assumption is corroborated by the fact that band ratio between near infrared band and red band, correlation is considerably poorer than that between other spectral bands (in certain cases, the correlation coefficient may be 0.5 or even less). All two-band indices are based on the simple physics: plants reflect less red light but more NIR radiation compared with non-vegetated surfaces. However, different indices have different advantages in retrieving plant information. A modified simple ratio (MSR) index is proposed for a better expectation of sensitivity to leaf biophysical parameters using remote sensing data. This index is formulated based on the evaluation of several two-band (NIR and Red) [1, 2].

Despite its use of the some indices saturates in cases of dense and multi-layered canopy and shows a non-linear relationship with biophysical parameters such as green LAI. Therefore, improved indices like the Renormalized Difference Vegetation Index (RDVI) and the Modified Simple Ratio (MSR) have been developed in order to linearity their relationships with plants biophysical variables. The MSR index produces images with a good contrast which could aid in making a reliable mapping of the plants cover of the area under study. This index was suggested as an improvement over RDVI in terms of sensitivity to plants biophysical parameter. The RDVI is less sensitive to the variations in the unknown foliage geometrical and optical properties and also less affected by the solar and view geometry. The MSR index is developed based on RDVI [3, 4, 1].

A broad histogram of the MSR image has a good contrast, which may help in detecting features with the ratio of standard deviation to mean (σ/μ). A narrow MSR image histogram with a small ratio may be broadened by histogram stretching. The histogram stretching should be avoided since it distorts the original values of the pixels, therefore a good criterion for the efficiency of a MSR image may be the ratio (σ/μ) [4].

The skewness characterizes the degree of asymmetry of a histogram distribution around its mean. Positive skewness indicates a distribution with an asymmetric tail extending towards more positive values. Negative skewness indicates a distribution with an asymmetric tail extending towards more negative values. Normal histogram distributions produce a skewness statistic of about zero [5].

The Hyperion sensor is the first hyperspectral imager on-board NASA's earth Observing-1 (EO-1) satellite that was launched on 21 November 2000. The EO-1 satellite follows the same orbit as Landsat-7 by about one minute. The spatial resolution of Hyperion is (30 m) and standard scene is (7.7 km) wide and (42 km) long. This sensor has (242) spectral bands ranging from (400 to 2500) nm, recorded at 12-bit radiometric resolution [6].

Spectral channels from 1-70 are collected from the VNIR and channels 71-242 are collected from the SWIR for each pixel location. The radiance values determined within the Hyperion bands, the SWIR bands have a scaling factor of (80) and the VNIR bands have a scaling factor of (40) applied [7, 8].

VNIR L = Digital Number / 40

SWIR L = Digital Number / 80

Hyperion's structure is the best model that contains data showing strength, stability and large surface area for living. It is a database can be used to develop and allows users the flexibility to view all information in different formats and groupings [9].

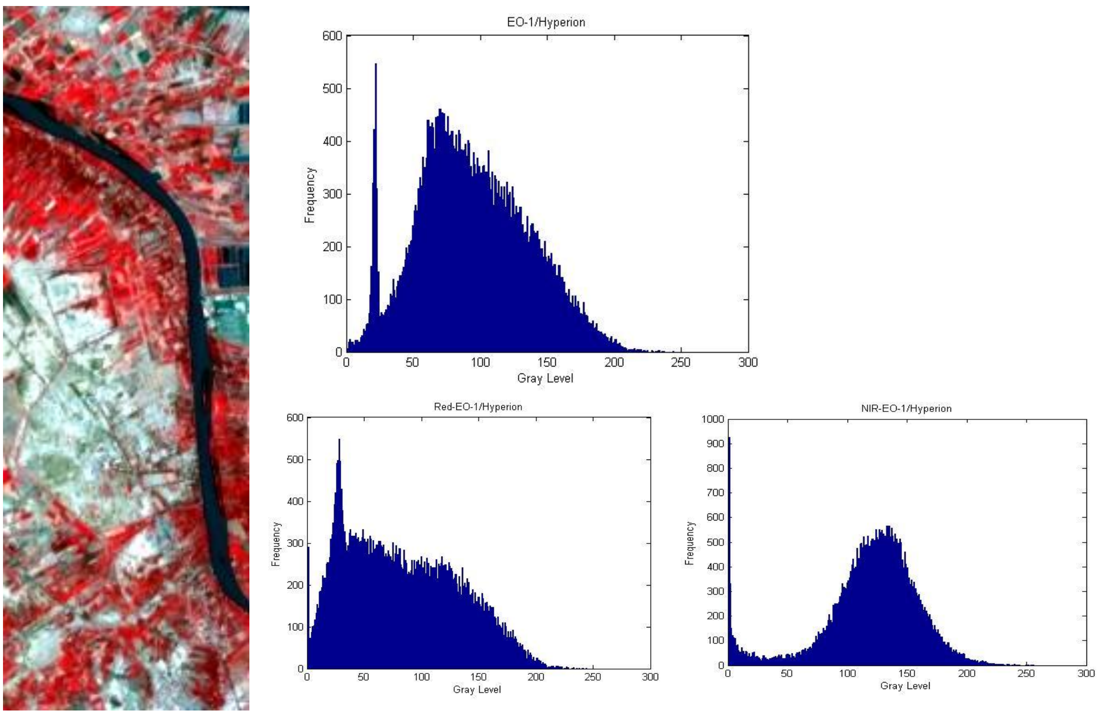
Generally, this study is comprised of two important parts. The first involved MSR index was a better sensitivity to detect the plants from soil reflectance and more linearity by focusing on the behavior of the MSR image histogram (marginal probability density function G(MSR)). The second was skewness statistic computing. In the present research, the re-verification of the G(MSR) derivation for the MSR index was performed. All these methods were implemented using ArcGIS9.3, ENVI 4.5 and MATLAB7.9b softwares. Figure-2 shows the overall flow of the steps that had been implemented in this study.

The Method and Material

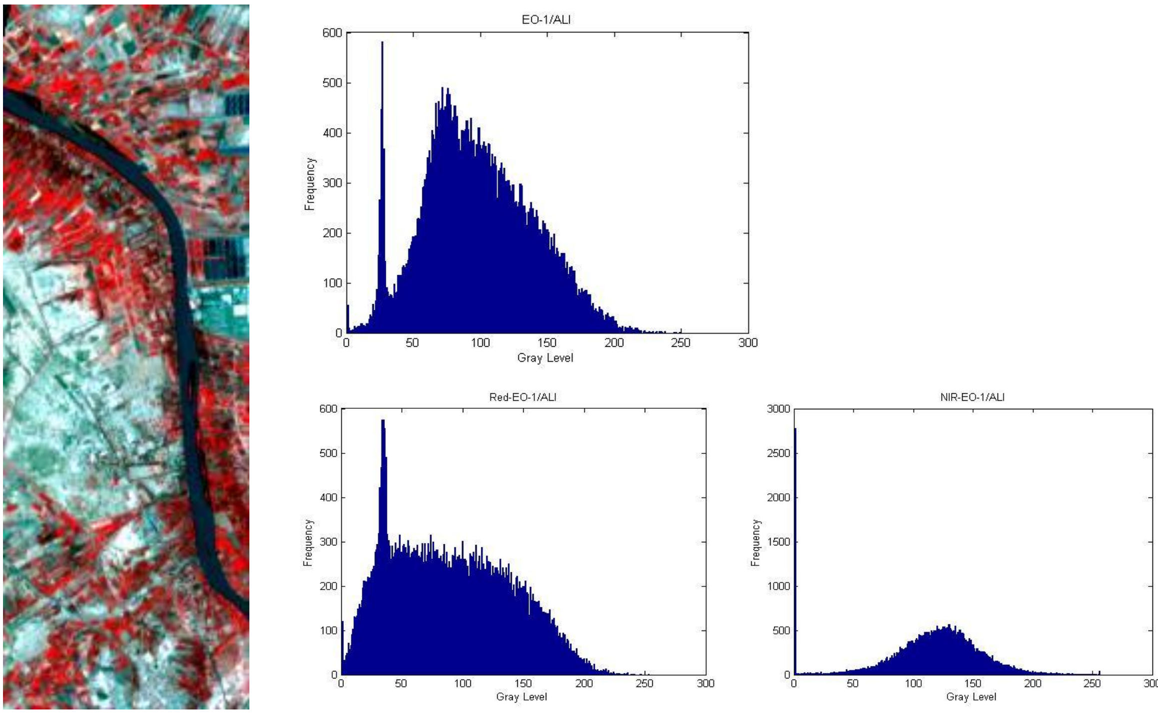
1- Research region and its corresponding data

Jurf Al- Sakhar quarter is located about 60 km south-west of Baghdad, the northern city of Musayyib, 13 km in the province of Babylon. Inhabited by peasants which doing in Agriculture field, being located on the Euphrates River from the right side of the river, abound where the cultivation of palms and fruit trees, as well as the cultivation of field crops such as wheat and barley. Associated with Jurf Al- Sakhar geography flat open with the city of Fallujah in Anbar province, with many areas of Anbar and form the eastern flank to district Fallujah and also relate to the cities of Latifiyah, Alexandria, Yusufiyah and Mahmudiyah.

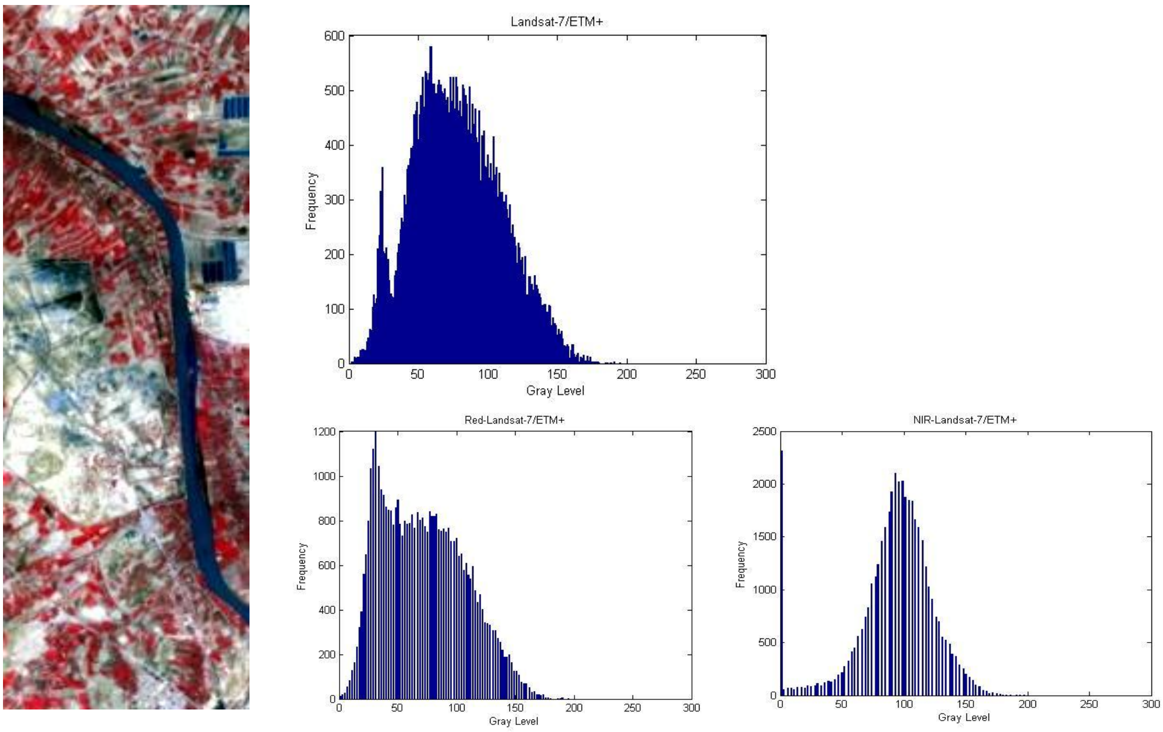
The area of the current research region about is (35, 17 Km^2). In this research, three satellite scenes with different sensors were used the Hyperion, ALI and ETM+. These scenes were obtained by Earth Observer-1 (EO-1) and Landsat-7 satellites exposure at 24 of September 2002. These scenes were previewed with three bands combination (R: 48, G: 31, and B: 20), (R: 6, G: 5, and B: 4) and (R: 4, G: 3, and B: 2) for Hyperion, ALI and ETM+ data respectively, which represent as three bands combination (NIR, red and green) bands, with the same spatial resolution of (30 m) for all these spectral bands. The available data for this region are shown in figure -1.



EO-1/Hyperion Scene



EO-1/ALI Scene



Landsat-7/ETM+

Figure 1- Original scenes (EO-1/Hyperion, EO-1/ALI and Landsat-7/ETM+) with its histograms, Red & NIR histograms

2- Modified soil ratio (MSR) index

Modified soil ratio (MSR) is a slope based vegetation indices group comprises simple arithmetic combinations of reflectance measurements, contrasting the high infrared and low red reflectance that characterizes photosynthetic vegetation. This contrast has been used widely in many applications which include the monitoring of vegetation cover. Pixel values in this group produce vectors with different slopes though the origin of the Red and NIR bi-spectral plot [6].

This index based on the repressing the effects the variety of soil reflectance with better sensitivity as improving the linearity, and this is developed based on Renormalized Difference Vegetation Index (RDVI) index by the quantified equations of MSR and RDVI are the following [1]:

$$MSR = \frac{RDVI}{\sqrt{NIR}} \text{------------------------------------} (1)$$

$$RDVI = \frac{NIR - Red}{\sqrt{NIR + Red}} \text{------------------------------------} (2)$$

$$MSR = \frac{\frac{NIR}{Red} - 1}{\sqrt{\frac{NIR}{Red} + 1}} \text{------------------------------------} (3)$$

The (RDVI) index was smaller than MSR, because the Red band was increased. The ratio NIR to Red bands is given by the following mathematical expression, and then compensates in the equation- 2 [4]:

$$r = \frac{NIR}{Red}$$

$$MSR = \frac{r - 1}{\sqrt{r + 1}} \text{------------------------------------} (4)$$

The algorithm of MSR index is stated in the followings steps:

1. Employ the equation-4 of the MSR index on three scenes data in figure-1, as shown in figure-2.

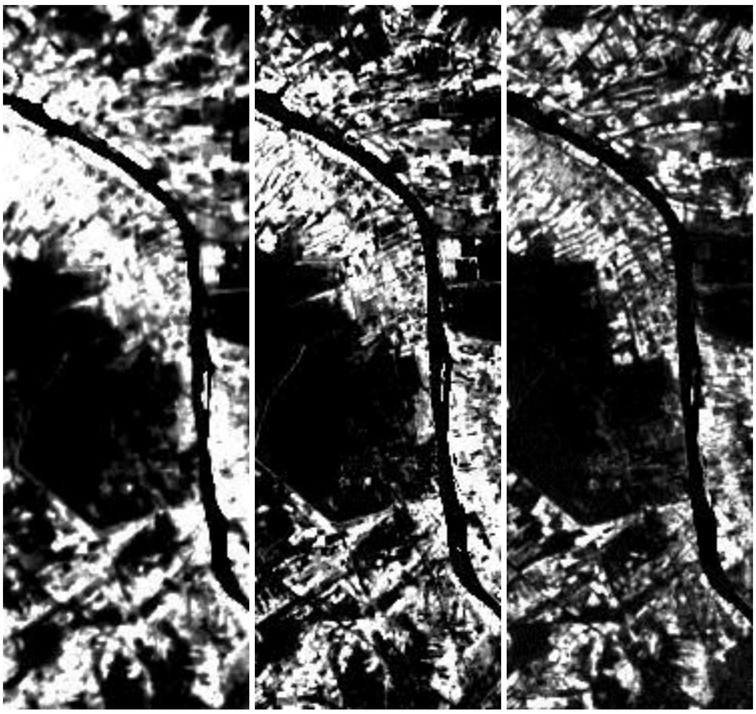


Figure 2- MSR images for three satellite scenes

2. Compute the mean (μ), standard deviation (σ) and the ratio (σ/μ) on MSR image, as the observed values of MSR data.

3. Draw the observed MSR image histogram.

4. Extract the (r) value as indicative of MSR value, as follows:

$$r(MSR) = \frac{MSR\sqrt{8 + (MSR)^2} + 2 + (MSR)^2}{2} \text{------------------------} (5)$$

5. Calculate the ratio square of the standard deviation of Red to standard deviation of NIR (λ) and the analytical expression *f* [r(MSR)] function of r(MSR) values in equation- 5 by the equations- 6 & 7, [4].

$\lambda = (\frac{\text{stdv (Red)}}{\text{stdv (NIR)}})^2$--- (6)

$f[\text{r(MSR)}] = \frac{2\lambda(\text{r(MSR)})}{(\lambda(\text{r(MSR)})^2 + 1)^2}$--- (7)

6. Determine the marginal probability density function (G) on MSR data using analytical expression f [r(MSR)] function by the equations- 7 & 8 [4], to extract the final formula of G(MSR) in equation- 9:

$\text{G(MSR)} = f[\text{r(MSR)}].\left|\frac{\text{dr}}{\text{d(MSR)}}\right|$---(8)

$$\text{G(MSR)} = \frac{\lambda\left((\text{MSR})\sqrt{8+(\text{MSR})^2} + 2 + (\text{MSR})^2\right)\left((\text{MSR})\sqrt{8+(\text{MSR})^2} + 4 + (\text{MSR})^2\right)}{\sqrt{8+(\text{MSR})^2}\left[\frac{\lambda}{2}\left((\text{MSR})\sqrt{8+(\text{MSR})^2} + 2 + (\text{MSR})^2\right)^2 + 1\right]^2} \text{-----(9)}$$

7. Compute the mean (μ), standard deviation (σ) and the ratio(σ/μ) for G(MSR) data, as the theoretical values of MSR data.
8. Draw the (G(MSR)) histogram.

The magnitude of the theoretical values of the image histogram may be useful in assessing the performance of the MSR index. If they are high, the histogram is broad, the image is good contrast and different land cover types are expected to be more clearly expressed. The low theoretical values mean a not good tonality contrast, which may imply difficulties in recognizing features of region interest.

3- Skewness approach

The statistical skewness measure is an important and widely used as the new tool to detect a small land cover area by extracting the detail information. In addition to feasibility of skewness in enhancement the edges of features in the images. The Mathematicians formula expressing of skewness is [10]:

Skewness = $g_1 \times (\frac{m_3}{m_2^{\frac{3}{2}}})$--(10)

Where:

$m_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2,\ m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3,\ g_1 = \frac{\sqrt{n(n-1)}}{n-2}$

n, x_i, $\bar{x}$ represent the number of all pixel in the scene, a set for all pixel values and the mean of gray scale values, respectively in a moving window.

m_2 and m_3, represent the second and third moments around the mean value, also the m_2 represent the variance value.

When dealing with the discrete values distribution, the estimator of skewness may be undefined (0/0) without using the factor (g_1) [11].

The steps of skewness algorithm are stated as the followings points:

1. Select the suitable window size is (7×7) can be exposed the symmetrical features information and it is the best for computations.
2. Apply the skewness approach, which defined by equation- 10 on MSR data in figure- 2, as seen in figure- 3.

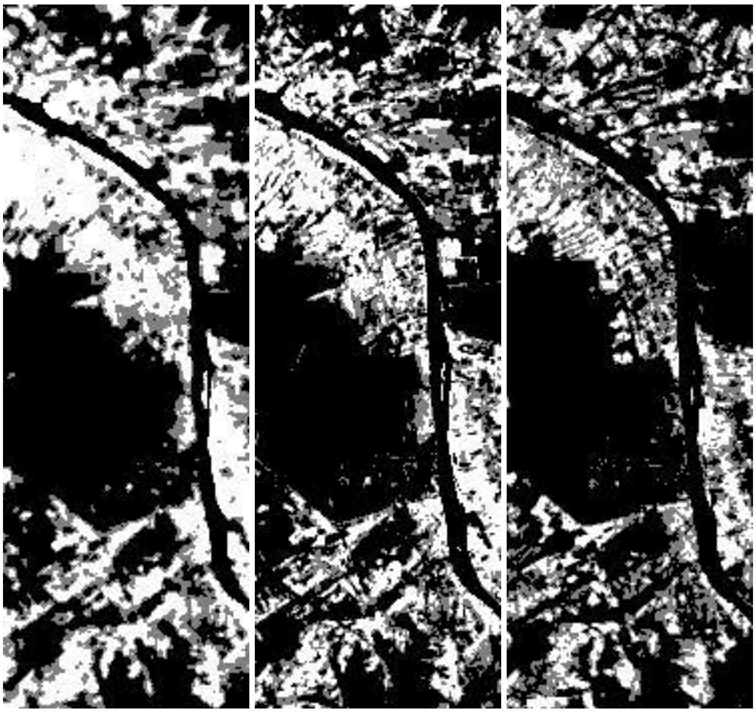


Figure 3- Skewness images for MSR images results

3. Combine the skewness results with Red and NIR bands of the original scene for classification, in three bands combination (R: MSR skewness image, G: NIR, B: Red). The classified images were generated as figure-4.

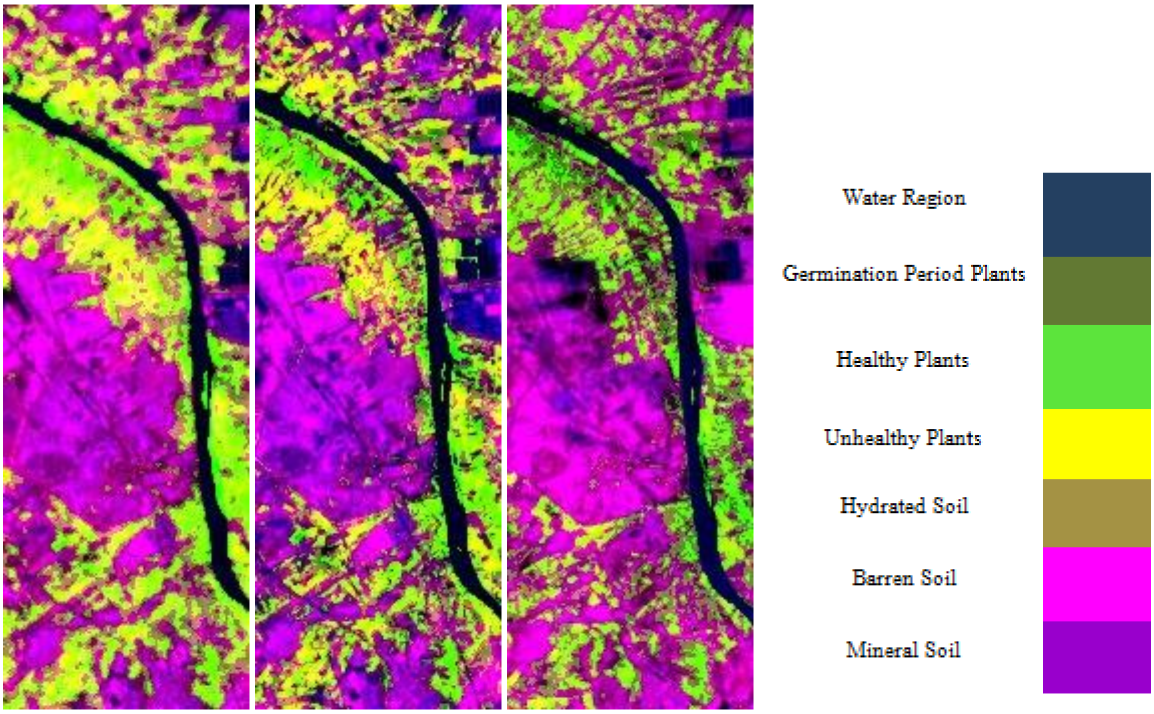


Figure 4- Classified images using skewness.

Results and Discussion

In this approach, the MSR index was performed of detection the plants classes using statistical estimating and mathematical form to evaluate MSR index efficiency. Depending on the correlations between bands which they should be taken into account.

The theoretical and observed calculations for MSR data were compared by various sensors, in order to the validity of the proposed methodological approach, which useful in classification the land cover regions.

Our results of calculation the equation-1 for MSR index on the three satellite scenes presented on Figure- 2. The bright white, bright gray and gray colors in these results indicate to the healthy, arid and germination period plants respectively. The reflectance of hydrated soil affected on the reflectance of germination period plants. These features can be separated and distinguished using the skewness technique as offered in figure- 3. This technique was as a threshold process. Figure-4 demonstrates the classification results by merging the results of figure-3 with the NIR and Red bands of the original scene to enhance the separating, because the pure multispectral bands are proper for extract the small area of the variety features (e.g., crops, palms and fruit trees).

Figure-5 displays the stretching of histogram distribution of the observed MSR values, when the marginal probability density function (G) implemented on this values for the three scenes, and they rescaled values for ranging between (-1 to ∞). The G(MSR) function is the histogram distribution of theoretical MSR index. Therefore, they are important to obtain broad histograms of MSR index for these scenes without any processing.

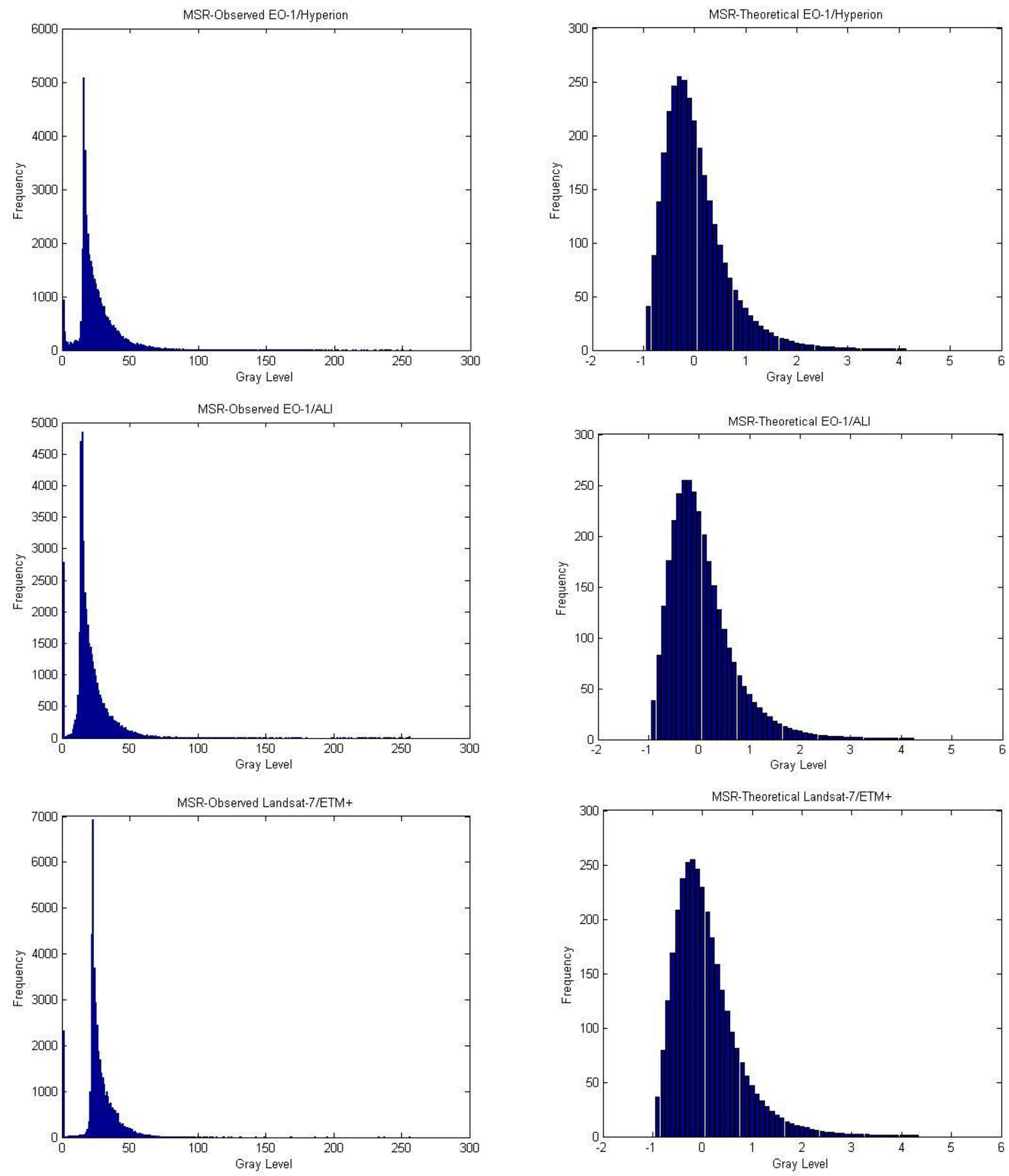
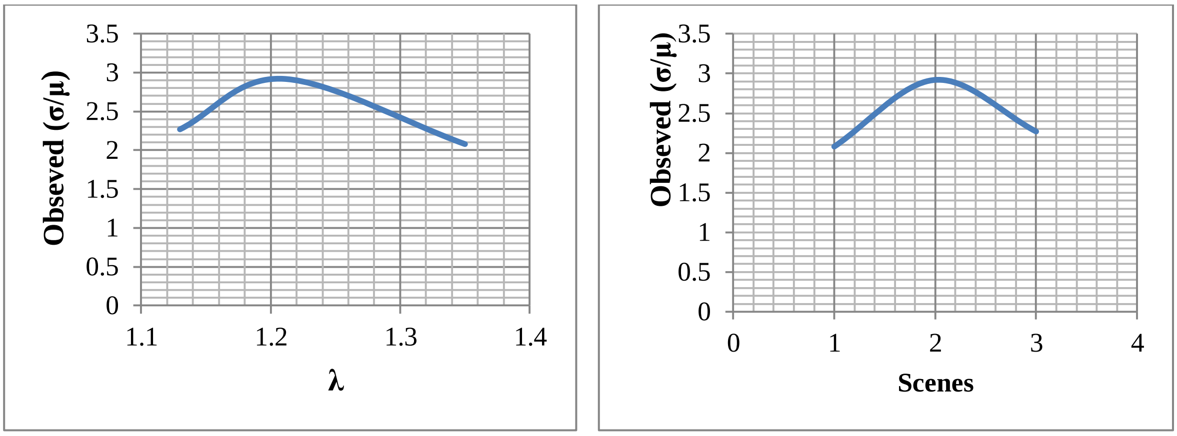


Figure 5- The histograms of the observed & theoretical (G(MSR)) MSR values for three scenes respectively

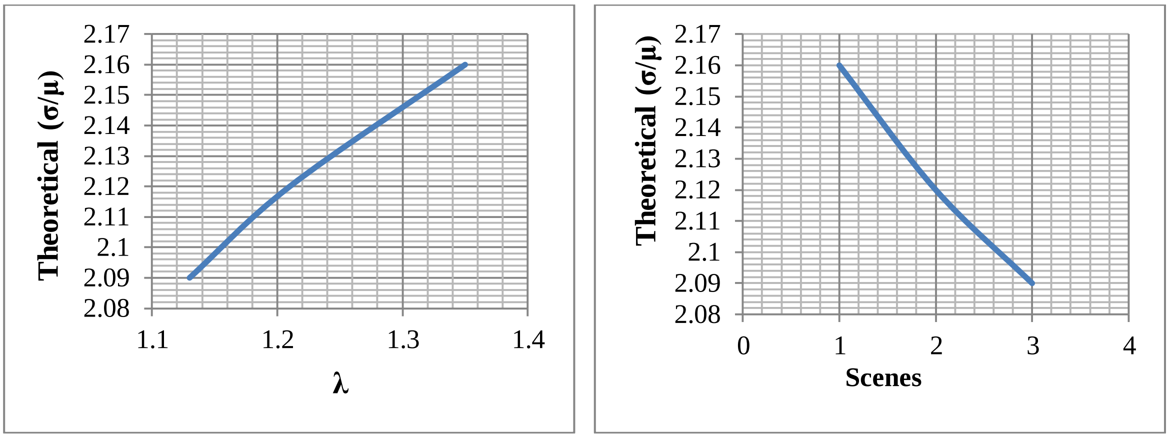
The MSR index is a non-linear index due to its non-linear relationship between the NIR and Red bands, as seen from alteration of observed ratio (σ/μ) values of MSR data with (λ) in figure-6 (a) and original scenes data in figure-6 (b).



a b

Figure 6- Graphical representation of the; (a & b)- variation of observed (σ/μ) values with (λ) and scenes

The alteration of theoretical ratio (σ/μ) values of G(MSR) data with (λ) values and original scenes data show that the MSR index has a linear behavior, when the function marginal probability density function (G) applied it, as displayed in figure-7 (a & b) respectively.



a b

Figure 7- Graphical representation of the; (a & b)- variation of theoretical (σ/μ) values with (λ) and scenes

The theoretical ratio (σ/μ) values were listed in table- 1 for MSR images were a good agreement and larger then observed ratio (σ/μ) values according to chart of figure (3b), for (λ=1.35 and λ=1.13) of the Hyperion and ETM+ scenes, except ALI scene was fewer ratio (σ/μ) values for (λ=1.21).

Table 1- Observed and theoretical statistical calculations of a MSR images

Scene	Mean (μ)	Standard deviation (σ)	Ratio (σ/μ)
Observed calculations			
EO-1/Hyperion	6.17	8.18	1.33
EO-1/ALI	4.58	7.23	1.58
Landsat-7/ETM+	1.57	1.86	1.18
Theoretical calculations			
EO-1/Hyperion	50.49	78.69	1.56
EO-1/ALI	52.33	79.84	1.53
Landsat-7/ETM+	53.23	79.97	1.50

Figure-8 offers the distributions of G(MSR) histograms, according to the (λ) values in table-2 for the scenes sensors. These positive skewness distributions prepare easily understood impressions of these scenes. Hyperion scene has a large value of (λ), because this scene can provide with statistical parameters (μ and σ) accurate quantitative estimation of the plants classes as explained from the classification images in figure- 4.

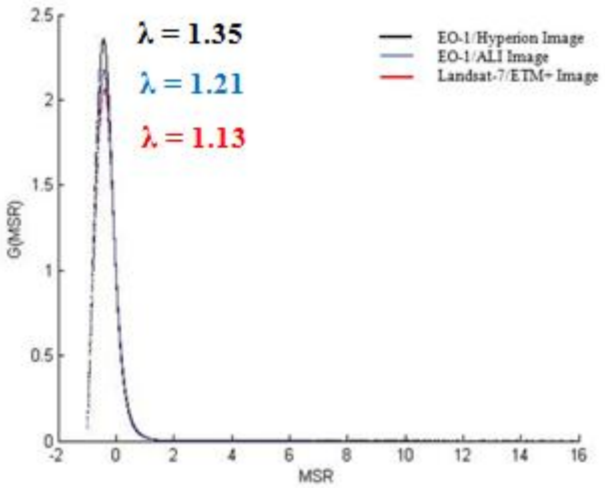


Figure 8- brightness values distribution of G (MSR) for three scenes.

Table 2- The ratio of σ(Red) with σ(NIR) values of the three scenes

Scene	λ
EO-1/Hyperion	1.35
EO-1/ALI	1.21
Landsat-7/ETM+	1.13

Conclusions

A novel algorithm has been introduced for distinguish and quick separation between features reflection in study area. This method adopted MSR index with skewness statistical based on Hyperion, ALI and ETM+ intensity information.

The MSR can be employed on original input feature space for scenes to solve a linear problem in that space. In this research, the MSR index linearity was proved by computation the marginal probability density function (G(MSR)) for MSR values (theoretical value) as shown in figure-7 (a & b), in order to good contrast, reduce and eliminate the influences of the variety soil reflectance underneath plants (such as; mineral and clay soils). Thus helps to best possible separation of the features, could be work in future. MSR index was made as a tool to improve the linearity. In figure-6 (a & b) the observed MSR value is a non-linear index due to its combination the band ratio (NIR/Red) with the different non-linear behavior for these bands, as displayed of the histogram distribution, for these bands in figure -1,. This index does on the classification of cases and the types of plants only.

The validity of G(MSR) equation for MSR index was re-derived in this work as shown in equation-9. This derivation may be applied to all indices in other fields, such as mineral soils exploration, that include the band ratio between two spectral bands.

Skewness technique applied on MSR values to exploit the detail information when combined with multispectral bands as unsupervised classification. This is a new and an effectively method for extensively observing the types of features of different sensors and its changes. Especially, between the crops and palms, fruit trees areas.

This study region characterized by the cultivation of crops from vegetables, fruit and palm trees. This technique was as a tool to determine of threshold process in this paper. The methodology adopted in the present paper may be used efficiently for other indices.

References

1. Chen, J. M. **1996**. Evaluation of vegetation indices and modified simple ratio for boreal applications. *Can. J. Remote Sens*., 22, pp: 229–242.
2. Vaiopoulos D., Skianis G. A and Nikolakopoulos K. **2004**. The contribution of probability theory in assessing the efficiency of two frequently used vegetation indices. *Int. J. remote sensing, Taylor & Francis*, 25(20), pp: 4219–4236.
3. Haboudanea D., Miller J. R., Pattey E., Tejadad P. J. and Strachan I. B. **2004**. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Elsevier Inc*. Remote Sensing of Environment 90, pp: 337–352.
4. Aim S. G., Vaiopoulos D. and Nikolakopoulos K. **2007**. A study of the performance of the MSR vegetation index, using probabilistic and geostatistical method. *Bulletin of the Geological Society of Greece*, Proceedings of the 111h International Congress, Athens.
5. Brown J. D. **1997**. Skewness and kurtosis. *JALT Testing & Evaluation SIG Newsletter*, 1(1), pp: 20-23.
6. Jafari R. **2007**. Arid Land Condition Assessment and Monitoring using Multispectral and Hyperspectral Imagery. PhD. Thesis. Discipline of Soil and Land System, School of Earth and Environmental Sciences, the University of Adelaide, Australia.
7. Barry P., Segal C. and Carman S. **2001**. EO-1/ Hyperion Science Data User’s Guide. TRW Space, Defense & Information Systems.
8. Beck R. **2003**. EO-1 User Guide v. 2.3. Cincinnati.
9. Hyperion space settlement. 2011.
10. Doane D. P. **2011**. Measuring Skewness: A Forgotten Statistic. *Journal of Statistics Education*, 19(2).
11. Joanes, D. N. and Gill, C. A. **1998**. Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society (Series D)*, 47(1), pp: 183–189.