REPUBLIC OF TURKEY
FIRAT UNIVERSITY
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

SIMILAR IMAGE RETRIEVAL IN ELECTRONIC COMMERCE FOR ONLINE SHOPPING BASED ON COLOR AND EDGE DIRECTIVITY DESCRIPTOR

Soran Abdulkarim Pasha
(151137110)

Master Thesis
Department: Software Engineering
Supervisor: Asst. Prof. Dr. Cafer BAL

ELAZIĞ

REPUBLIC OF TURKEY
FIRAT UNIVERSITY

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

SIMILAR IMAGE RETRIEVAL IN ELECTRONIC COMMERCE FOR ONLINE SHOPPING BASED ON COLOR AND EDGE DIRECTIVITY DESCRIPTOR

MASTER THESIS

Soran Abdulkarim Pasha
(151137110)

Department:Software Engineering

Thesis Submitted Date: 09 May 2017

Thesis Defense Date: 24 May 2017

Thesis Supervisor: Asst. Prof. Dr. Cafer BAL (Firat U.)

Other juries: Assoc.Prof. Dr. Muhammed Fatih TALU (İnönü U.)

Assoc.Prof. Dr. Bilal Alataş (Firat U.)

May – 2017

DECLARATION

I am presenting this thesis with title “Similar Image Retrieval in Electronic Commerce for Online Shopping Based on Color and Edge Directivity Descriptor” for the requirement of Master’s degree in Software Engineering at Firat University. I declare that proposed system in thesis is my own work with all simulations and programming and has not been submitted for the award of any other degree at any institution.

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who raised and taught me that even the biggest task can be accomplished if it is done patiently and one step at a time. I would like to thank the rest of my family members and my wife for their understanding, moral supports, encouragements, prayers, patience and all kind of support.

ACKNOWLEDGEMENTS

First of all, I gratefully acknowledge the support I received from several people which have helped me in my study. I am indeed very fortunate to have such an affectionate bunch of friends and well-wishers. I owe them all many thanks. It will be a mistake on my part not to mention some of their names here to whom I extend my heartfelt gratitude. I sincerely thank:

Supervisor: Asst. Prof. Dr. Cafer BAL, my guide, for being so kind, caring and generous and for devoting so much time for me, which I do not deserve, in spite of his terribly busy schedule. I have many special thanks to him.

Very special thanks to my family who always kept on urging me to concentrate on my studies and worry about nothing even when things at home were far from being fine. Finally, I appreciate the role of Firat University and Faculty of Technology Software Engineering department for giving me this great chance to study and got a certificate that will never be forgotten. Hope you all the best and delight. Wish all the best to all. Also, I want to thanks my wife who supported and helped me every time.

All my lovely friends for continuing to be a source of inspiration, and for all those precious moments which gave me a sense of direction even when I was utterly helpless. And of course, for their invaluable collection of books to which I am yet to return quite a few books.

Sincerely

SORAN ABDULKARIM PASHA

Elazig - 2017

LIST OF CONTENTS

Page No

ABSTRACT

SIMILAR IMAGE RETRIEVAL IN ELECTRONIC COMMERCE FOR ONLINE SHOPPING BASED ON COLOR AND EDGE DIRECTIVITY DESCRIPTOR

E-commerce is attractive a common choice for buyers. Actually, the popular item searching method that e-commerce websites give is keyword search. And the consumers should be accurate choose relevant keywords to search for items. This thesis presents a method based on similar image retrieval in e-commerce for online shopping based on color and edge; aiming at efficient retrieval of images from the large database for online shopping. Here, RGB (horizontal and vertical) projection is used for creating our application with a huge image database, which compares image source with the destination components. This method is proven to be one of the best techniques for online shopping product search on the Internet. In e-commerce business transactions, buying and selling products are made through the electronic system or via the Internet. In this thesis, a technique is used for finding products by image search, which is convenient for buyers in order to allow them to see the products. The reason for using image search for products instead of text searches is that products searching by keywords or text have some issues such as errors in search items, expansion in search and inaccuracy in search results. This technology is providing a new search mode, searching by image, which will help buyers for finding the same or similar image retrieval in the database store. The image searching results have been made customers buy products quickly. The results of the implementation show that searching process for products in e-commerce different between search by image and search using text for buyer option.

Keywords — color, E-commerce, image search, retrieval, online shopping, product, technique.

ÖZET

ELEKTRONİK TİCARETTE ÇEVRİMİÇİ ALIŞVERİŞ İÇİN RENK VE KENAR YÖNELİM AÇIKLAYICI TABANLI BENZER GÖRÜNTÜ ERİŞİMİ

E-ticaret, alıcılar için cazip bir seçimdir. Aslında e-ticaret web sitelerinde anahtar kelime yöntemi ile nesne araması yapılır. Tüketicilerin, aranan nesneye ilişkin anahtar kelimeyi kesin olarak seçmesi gerekir. Bu tez, e-ticarette çevrimiçi alışveriş için, büyük bir veri tabanında verimli bir şekilde resim erişimini amaçlayan, renk ve kenar tabanlı benzer resim erişimine dayalı bir metot sunmuştur. Burada büyük bir veri tabanı ile uygulamanın oluşturulması için, resmin kaynak ve hedef bileşenlerini karşılaştıran, RGB (yatay ve dikey) iz düşüm kullanılmıştır. Bu metodun internet üzerinden çevrimiçi alışveriş için en iyi ürün arama tekniklerinden biri olduğu kanıtlanmıştır. E-ticarette ürünlerin ticari işlemleri, satın alma ve satışı elektronik sistem sayesinde veya internet aracılığıyla yapılır. Bu tezde resim arama ile ürünleri bulmak için, alıcılara uygun ürünleri görmelerine izin veren, bir teknik kullanılmıştır. Ürün arama için metin arama yerine resim kullanmanın sebebi, anahtar sözcük veya metin ile arama yapmanın, hatalı ürün arama, geniş ölçekte arama ve doğru olmayan arama sonuçları gibi, bazı sorunlarının olmasıdır. Bu teknoloji, veri tabanından benzer veya ayı görüntü erişimi için alıcılara resim aramada yardımcı olacak yeni bir arama şekli sağlamaktadır. Yapılan arama sonuçları ile tüketici ürünü çabucak satın alabilir. Uygulama sonuçları alıcı için e-ticaretteki ürün arama sürecinde, metin ve görüntü arama arasındaki farkı göstermiştir.

Anahtar Kelimeler — renk, E-ticaret, görüntü arama, geri getirme, çevrimiçi alışveriş, ürün, teknik.

LIST OF FIGURES

Page No

LIST OF TABLES

Page No

LIST OF ABBREVIATIONS

E-Commerce : Electronic Commerce
RGB : Red, Green and Blue
IR : Image Retrieval
CBIR : Query Based Image Content
MPEG : Moving Picture Experts Group
E-Business : Electronic Business
B2B : Business-to-Business
B2C : Business-to- Consumer
B2G : Business-to- Government
C2C : Consumer-to-Consumer
AI : Artificial Intelligence
ASP : Active Server Page
SQL : Structure Query Language
GUI : Graphical User Interface
SSIM : Structural Similarity
MVC : Model, View and Controller
USD : United States Dollars
PHP : Personal Home Page
CSS : Cascading Style Sheets
NSA : Negative Selection Algorithm
DPS : Direct Pixel Similarity
LCH : Local Color Histogram
GCH : Global Color Histogram
SIFT : Scale-Invariant Feature Transform
PRESS : Proposed Polar Raster Edge Sampling Signature

1. INTRODUCTION

The development of Internet technology has eased the increase of in-home shopping [1]. The Internet has had a powerful effect on marketing and created another type of retail transaction called e-commerce for online shopping. Electronic commerce (e-commerce) is a term used for selling and buying on the Internet. Moreover, e-commerce contains several groups in the same platform, for example, online funds transfer, Internet marketing, electronic data interchange, inventory network management, online transaction preparation, and data collection. Businesses can get orders, sell products, and get payments through the web [2].

For this reason, nowadays, many businesses have begun building business websites to expand the selling and buying of products everywhere in the world. There are some problems in online shopping transactions, such as it is sometimes hard to guarantee security or protection on site exchanges due to the lack of trust and the absence of system safety, resulting in poor execution of e-commerce. It may be hard to combine e-commerce systems or sites with a database. This thesis is an attempt to solve the main problems, especially security and accuracy of image retrieval in an e-commerce database. In this thesis, a good database has been created for data collection.

This thesis is deploying a technique that uses low-level features, which are extracted automatically from images and then are used for indexing and retrieval. It combines color information in a histogram and with e-commerce for online shopping. The idea of this work is using image processing to aid in shopping online. Our system permits the buyer to upload an image and then return comparable products using image retrieval systems [3].

Retrieval of images in e-commerce is normally managed through a manual keyword insertion defining a scene. Especially, the search primarily depends on the keyword search [4]. In addition, image retrieval is increasing in many fields, for example, medical, private life, modern/commercial, product, medicine, and workmanship [5]. In addition, e-commerce is an increasing power in the creation economy, while the web has

become the main source of trade products, e-commerce services, and information for many people worldwide [6, 8]. Access to Internet-related information has become a daily activity for many people in the digital century [9].

Recently, expanding accessibility of vast multimedia information, fast development of images in huge databases and customers in different areas require successful and productive image retrieval frameworks for dealing with visual data [10]. Online shopping has turned into an effortless process. With various distinctive and insightful advances improving the work, buyers require performance to locate an impeccable match. In addition, this technology is a smart search engine that cannot just retrieve several bits of product images by its text or color, and the machine itself does not comprehend the way people do – it just retrieves the results by comparing tags determined by people [11].

Categorizing products exactly and efficiently is a big objective in current e-commerce. The products are commonly described by metadata, such as image, title, and other aspects, with most assigned physically by the sellers. Products are uploaded to the e-commerce website and are naturally located in many categories. In addition, categorizing products helps e-commerce sites provide users the best shopping experience [12].

As shopping online is an accessible area worldwide and is continuously expanding with differ types of new mechanisms, more specifically connecting between image search engines and e-commerce, some literature reviews regarding other works have already have been discussed. The proposed mechanism can have variant uses, but we are now focusing on some parts of the implementation to try to connect them with local information of the same type.

Representing images with number values to describe their special visual properties and content is a difficult operation. A large combination of early planning and innovative methods of image retrieval may be found in well-structured studies, and authors have discussed important challenges to create systems that may be used in every place [13].

In [14], researchers gave an overview of a huge set of features for CBIR and compare them with other types of tasks and have been created a power database that are stored images and data and retrieval images of the features are analyzed in detail. While feature specific studies estimating color characterization and the real aim of these works, it was to get the best descriptors for retrieval images. [15, 17], texture description, in this work the researchers are increased interest in returning images in a huge gathering or from the database. All images have to be described or appeared by clear features [18, 20] and shape description [21, 22] strategies, clearly, outline the many ways researchers explored in the search of effectively representing image content. In [23], three content-based image sorting In [24], an efficient computer-aided retrieval of an image based on plant images (a leaf) was proposed using texture and shape features proposed for e-commerce, particularly in the medical industry. According to this research, the users or anyone can find plant images among many difference kinds of images from a large database.

In [25, 26], several studies have identified factors affecting Internet sales. Color and text features are used in the QBIC [27], SIMPLIcity [28] and MIRROR [29] IR search. The descriptors that are advised by MPEG-7 [30, 31] for indexing and retrieval keep equivalence among the size of the feature and the product of the results. In [32], authors used a content-based image retrieval system method created by adopting a strategy of joining various features of color, shape, and relevance feedback. This feedback is used as a helping feature to make the system more active and exact.

2. E-COMMERCE AND PROPERTIES

The term e-commerce is used for a broad area of trading activities on the Internet for products or goods [33]. It also relates to any kind of commercial transactions that connect parties electronically instead of directly [34]. Moreover, e-commerce often is linked with selling and buying online or making any transaction involving transfer of possessions or is used for service rights by the computer network [35].

The full definition of e-commerce is using e-communications and processing computerized information technology in trade transactions for creating, converting, and redefining relations for value creation among individuals and organizations [36]. In addition, e-commerce is challenging much of this traditional business contemplation [37, 38].

Image Search

Image retrieval is a specific information search used to retrieve images. To search for an image, a customer may give query terms, for example, keywords, image file/link, or tapping on an image, and the framework will return pictures ‘similar’ to the query [39]. Image retrieval is used for searching for products or items online by uploading an image, and it has improved the advancement of searching [40].

Normally, buyers use the search engine to search for products [41]. The visualization of searching by image satisfies customers with higher searching requirements, helps customers find product information more conveniently, and improves the online shopping experience.

2.1 Traditional Commerce VS E-Commerce

In Table 2.1 displays, the difference between traditional commerce and e-commerce is illustrated.

Table 2.1. Traditional E-Commerce vs. E-Commerce **[**42**]**

Traditional Commerce	E-Commerce
Difficult dependence on data exchanges from one to another.	Information sharing is facilitated by e-communication channels requiring little reliance on one-to-another data interchange.
Transactions finished in asynchronous ways. Manual mediation is used for every transaction.	Should be possible in an asynchronous way. The entire process is totally automated.
Hard to build up and maintain standard practices.	Standard practices should be effectively established and maintained in e-commerce.
Business communications depend on human abilities.	There is no need human mediation.
Depends on people for transaction.	E-commerce sites give the client a platform where all the
No uniform stage for data sharing, as it depends heavily on individual communication.	E-commerce gives an all-inclusive stage to support business/business activities over the globe.

2.2. The General Differences between E-Business and E-Commerce

There are discussions between specialists and academics on importance and restrictions imposed on both e-business and e-commerce [38]. In addition, e-commerce principally includes transactions that pass limits and the exchange of products among consumers, etc. Moreover, e-business basically includes the use of advanced technologies

to business forms inside of the enterprise, as shown in Figure 2.1 we think that success depends on the qualification between e-commerce and e-business [43, 44].

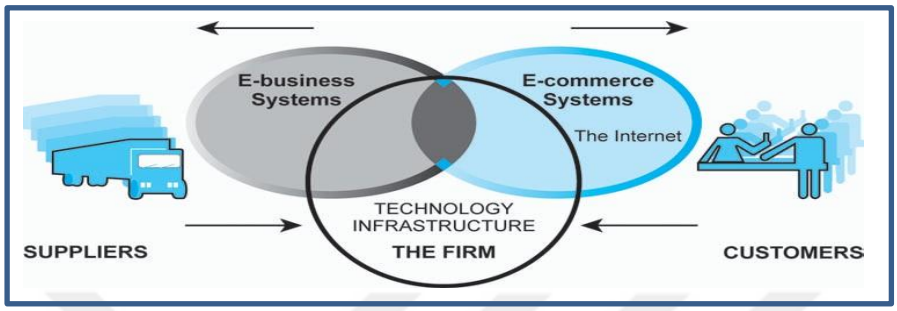

Figure 2.1. General Differences E-Business and E-Commerce [38]

2.3. The Different Types E-Commerce

2.3.1. B2B E-Commerce

Business to business (B2B) is also referred to as e-commerce between enterprises. This is the kind of e-commerce that deals with relations between businesses. Many people use this type of e-commerce, and most predict that B2B e-commerce may continue to develop more quickly than B2C. Figure 2.2 describes B2B e-commerce.

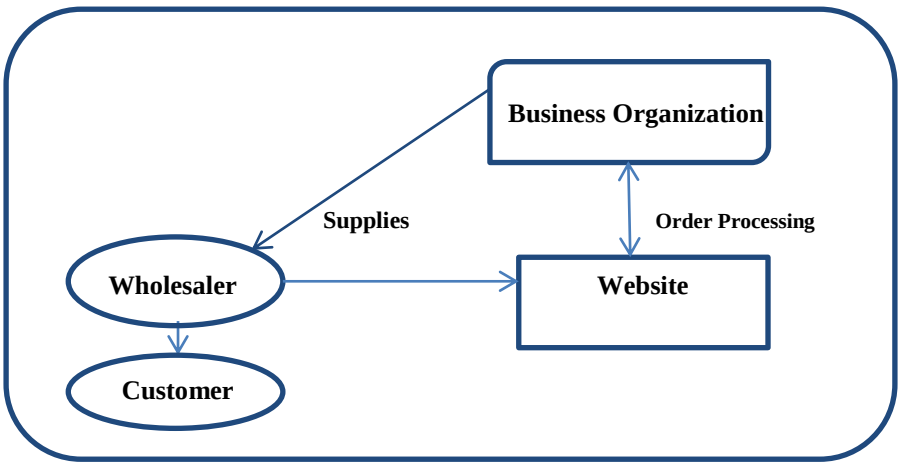

Figure 2.2 Business-to-Businesses [46]

The B2B market has two types: e-markets and e-infrastructure. Moreover, the e-infrastructure is the B2B structure, which basically includes the following [45]:

- Logistics – transport, storage, and delivery.
- Application service sponsors – publishing, hosting, and administration software packages from a central facility.
- Software content and simplification, online secured conduct, and delivery.
- Empowerment of commerce on the Internet.
- E-markets are known as the websites on which buyers and sellers interact with each other and negotiate deals [46, 47].

Using of B2B E-Commerce in Developing Markets

Transfer expenses, we have three costs. The 1st point is the reduction of the search of cost, as consumers, not important go over mediators to search for information about sellers, products and costs as in a customary of supply chains. In terms of the attempt, spent of money and time because the Internet is the largest skillful information channel. In B2B markets, customers and seller are gathering together in the same online trading center, decreasing costs of search in the far beyond.

The second is the decrease in the costs of handling contacts (methods of paying and receipts) because B2B provides automating of action process and because, the fast execution of the similar related with different stations. The 3rd one, online shopping procedure enhances stock organization and logistics [46]. B2B electronic markets and the suppliers are could react directly with consumers. For example, electronic markets are being taken into account as mediators because they through between suppliers and consumers in the supply chains [46]. Gathered for a great number of buyers and sellers in an individual electronic market detect market price information and deal operation for the share.

The Internet offers for the publication of information or transactions and makes the information easily and ready to all people of the e-market. High-cost clarity has the impact of destroying descending the cost differences in the market. In addition, B2B e-markets expand limits for the dynamic prices, while many buyers and sellers of shares in combined price fixing and directions auctions. In the e- marketplace needs the buyers and sellers, therefore aggregated to access cheap prices that are less than those changes resulting individual business [46].

2.3.2 B2C E-Commerce

Business to consumer refers to trade between consumers and companies, including customers gathering information, purchasing material products like concrete, or consuming commodity information (goods or digital content, like software or e-books), and receiving products via the electronic network [48]. This type is the most used after B2B e-commerce and is the simplest form of e-commerce. This origin could be attributed to the website retail trade [49]. The most famous B2C business models are online enterprises like Amazon.com [46].

In addition, B2C e-commerce reduces costs (search costs) by expanding consumer access to data that allows users to get the most competitive rates for products or services. Moreover, B2C e-commerce additionally reduces market entry barriers, where the cost of setting up and maintaining a website is less expensive than introducing architecture for a company. Furthermore, B2C e-commerce is appealing because it spares companies from figuring in the extra cost of the physical circulation system. In addition, for countries with a developing and powerful online populace, delivering information products is progressively feasible [46].

The main focus of B2C business models of e-commerce are on the products, which concern the management of personal private equity and financial affairs with electronic banking instruments [50]. Figure 2.3 explains the details about B2C model business.

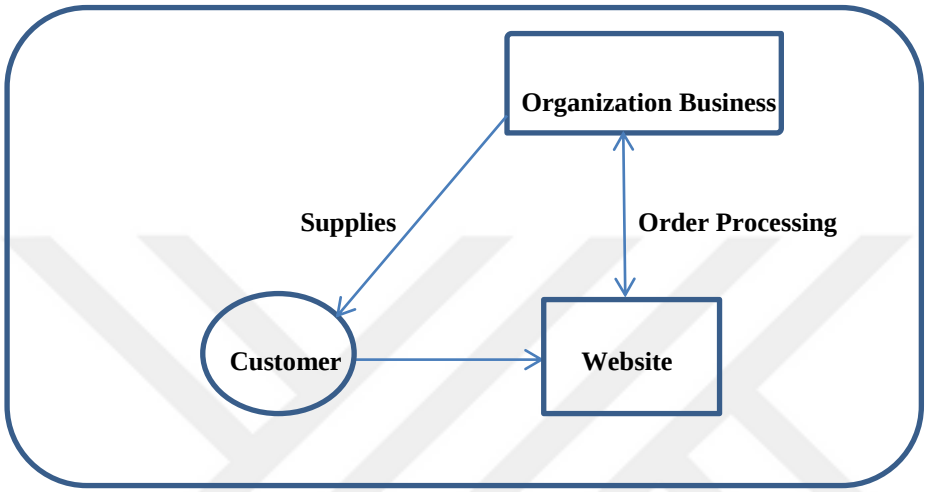

Figure 2.3. Business-to- Consumer [46]

2.3.3 B2G E-Commerce

Business to government (B2G) is commerce among companies and other sectors to exchange information with difference business organizations. As shown in Figure 2.4, this type of e-commerce has two properties:

- The public sectors assume an experimental/leadership role in the establishment of e-commerce.
- The other assumes that the public sectors have the most required to make the purchases system more efficient [51].

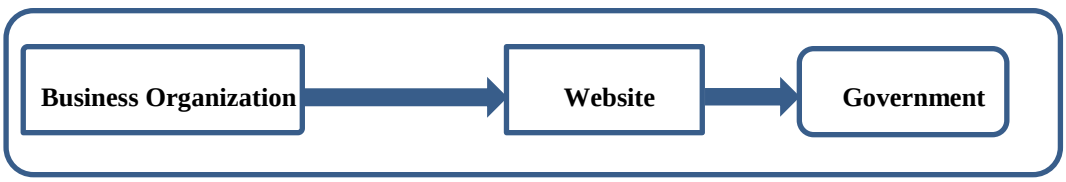

Figure 2.4 Business-to-Governments [46]

2.3.4 C2C E-Commerce

Consumer to consumer (C2C) is between private people or consumers [51], providing a new method to permit customers to interact with each other. This type of e-commerce is described by the growth of e-commercial centers and online buyers, especially in header sectors [52]. In any case, well-known C2C examples include eBay and Napster [46]. Figure 2.5 displays how to work this type of e-commerce.

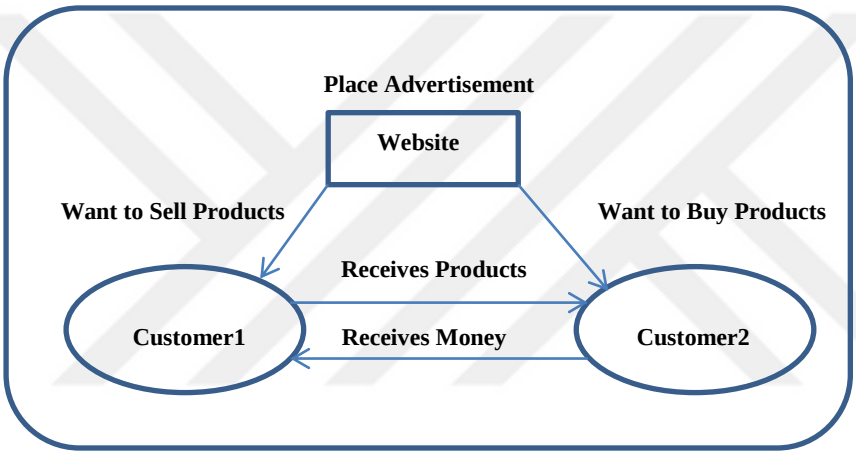

Figure 2.5 Consumer-to-Consumers [46]

2.4. Strategy Profitability for Enhancing E-commerce

Retail e-commerce profits from positive economic patterns and maintained buyer eagerness for online shopping [53]. We have some strategies that traders can use to improve service on the Internet and to open new markets and high cost.

2.4.1. Strategy 1: Expand into International Markets

International payment processors help take advantage of the possible limitless products in the market. According to an analysis of the United States online dealers by Internet Retailer, three-quarters of participants sell products internationally [54]. However, this same analysis demonstrates that online sellers support orders from buyers from the private business sector [54].

2.4.2. Strategy 2: Provide More Consumers More Ways to Pay

More choices in the kinds of payment choices decrease the number of deserted shopping carts. Payment choices can be a key factor when the buyer agrees to online shopping. How are online customers paying at checkout? What changes are happening in their payment choices? Answering these questions will open new doors for extending benefits that enhance the buyer's shopping experience [53].

2.4.3. Strategy 3: More Payment Options Means More Sales

Trading probably would not happen if the payment type had not been accessible. By understanding why these customers are not paying with traditional debit and credit cards, online merchants are in the best position to choose and provide new payment choices that may fulfill their buyers' needs and contribute to revenue development.

Why are online customers inclining toward other payment types? An important part of consumers doesn't have credit or debit cards while others do not want to increase credit card balances. Moreover, the most of the buyers have security worries about utilizing the credit or debit card online. In each shopping environment, buyers want to sure that place is an important to select, comfort, and security. Online customers need the flexible to choose ways to payment at checkout, pretty much as they do in their offline transactions [53].

Providing buyers with more payment options may also decrease shopping cart desertion costs. One of the reasons for leaving carts was not offering a favoured payment option [55]. There are several types of electronic payments:

- Credit card,
- Debit card, and
- Smart card [42].

2.4.4. Strategy 4: Leverage Business Intelligence

The payment operations produce valuable information to use for making smart business decisions [53]. One of the main advantages online businesses have over traditional brick-and-mortar stores is the capability to gather full information about what shoppers do when the buyers go to the sites [53].

2.5. Visual Search for Ecommerce Going Mainstream

Moreover, e-commerce without a search engine to build data makes little sense when we need difficult-to-find products. That is, the idea search has such significant effects for online shopping. The right way to do it provides service instead of products, which can only increase and grow [56]. There are three main reasons visual search experiences better acceptance, which are discussed next.

2.5.1. Increased Accuracy in the Same Time

Modern artificially intelligent systems are capable of producing exact results rapidly. This rapidity was missing in the previous systems.

2.5.2. Availability of Visual Search as a Service

The most popular e-commerce features increased larger acceptance after seller(s) made them easy to use as a service to sellers.

2.5.3. Advancements in Artificial Intelligence

Artificial intelligence (AI) is the capability of software to make decisions based on data supplied to it. Visual search is an AI challenge for software to identify the products in an image (uploaded by a buyer) and to complete a search to find similar products [56].

2.6. General Advantages and Disadvantages of Using E-Commerce

2.6.1. Advantages to Organizations

Using e-commerce can increase the market scale to global markets. An organization can simply locate more customers, the best suppliers, and the right trading share worldwide.

- E-commerce supports organizations by decreasing the cost of operation, distribution, recovery, and management of sheet-based information by digitizing the information.
- E-commerce improves the trademark image of the firm.
- E-commerce helps the organization provide the best service to customers.
- E-commerce helps streamline business operations faster and more efficiently.
- E-commerce decreases spread sheet work [57].

2.6.2. Advantages to Customers

- 24x7 supports: Customers can shop for products or services anywhere anytime. It means 24 hours a day, 7 days a week.
- E-commerce application availability gives members the most choices and the fastest delivery of items.
- E-commerce availability gives members the most choices to compare and choose the cheapest and best options.
- The customers can view or review comments about a product before making a purchase or can make comments after purchasing.
- E-commerce can provide default options for auctions.
- Customers can see the products before making a final decision to buy.
- It promotes competition between organizations and gives discounts to customers [57].

2.6.3. Advantages to Society

- The customer does not have to travel to shop; therefore, there is less traffic resulting in less air pollution.
- It helps people decrease the cost of products for people who are less wealthy.
- E-commerce can provide access to products to the countryside.
- In addition, it can help the government provide public services, such as social services, education, and health care at decreased costs and in the best way [57].

2.6.4. Technical Disadvantages

- There can be a lack of safety, reliability, and quality attributable to poor execution of e-commerce.
- Many times, it is difficult to combine e-commerce systems with databases, operating systems, or some other components.
- Server problems [57].

2.6.5. Non-Technical Disadvantages

- **User Resistance**: Sometimes buyers do not believe the sites ensure privacy, and sometimes, buyers may not be sure how to transact for products.
- **Physically**: The buyers cannot touch or feel products over on the Internet in online shopping [57].
- **Maintenance and Training**.

2.7. Developing in E-Commerce by Image Retrieval Engines

Since its appearance, e-commerce has developed in size. As an example, hundreds of online shopping sites are created in a year and more have developed over time. The Internet (e-commerce) includes several types of images and other visual information. A

search for images was developed more than 20 years ago [58]. The past systems were test-oriented or for industry-specific trained retrieval, like library digital images, helping in color retrieval and shape retrieval.

With the hasty development of technology, the Internet increased multimedia data on a large scale; the technology of image retrieval is rising quickly. For system ability, distributed design [59], MapReduce [60] multi-core technology, and architecture [61, 62] have all been announced to enhance the storage and efficiency of retrieval.

Retrieval of the image in e-commerce for online shopping is a developing part of research, which has a large commercial potential. Some firms turn to e-commerce-oriented IR, such as Hitachi's GazoPa [63].

Technically, product images commonly have clear image features and strong backgrounds, making it simple to compute feature operators and to make decision comparisons of similarity, highly improving retrieval accuracy. Sometimes, it can be used as a source of information (in e-commerce, the buyer wants to buy products while online shopping and uses the image for the search engine to help find a product [64].

2.8. How the Internet Changes Business

The Internet continues developing as a center for commerce, permitting an organization to conduct business anywhere anytime. Moreover e-commerce and the Internet eliminate imperatives of time and distance in commercial activity and empower a vast number of relations among shoppers, suppliers, and business partners [65]. The advantages of the new economy have expanded the level of competition in every commercial venture, and it offers a great chance for even poor companies to present new products or services as a result of the speed and decreases the cost of doing business. In today's new business environment, power has moved towards shoppers who request strong items that convey new measurements of cost (time and substance notwithstanding to the present measurements) cost and quality [66]. Electronic marketplaces also decrease

incompetence brought on by buyer search costs to get information about the cost and item offerings to add the cost of sellers to transfer information about their costs and item offerings [67, 68]. At the point at which a company interacts electronically with users, purchasing behavior may be analyzed so that the company can allocate its item and service performances to the single users in the new economy [69, 70].

The capability to customize products connected with the capability of sellers to get generous information about potential buyers, for example, demographics, preferences, and past shopping behavior, is a great opportunity to enhance sellers' capability of price separation that permits sellers to charge various costs for various buyers [71].

2.9. The Need for Image Retrieval from the E-Commerce

The Internet consists of an incredible number of images and other visual media, for example, recordings, motion, and images, that may have a place with organized collections (e.g., historical center accumulations) or be autonomous (e.g., images are found on the Internet, such as people's photographs and logos).

Instruments for active retrieval of this data can be exceptionally valuable for some applications. In the present area, we attempt to display why such tools are keys for the shopper, for which applications individuals require such tools [72].

2.10. The Effect of E-Commerce on the Developing Countries

Internet commerce will change the face of work until the end of time. In addition, e-commerce changes banking in the twenty-first century and has influenced the international economy in numerous ways. First, it has influenced information technology, and the economy. Moreover, online shopping has enhanced the efficiency of expansion.
A few countries are taking advantage of the results; they are presently in a position to benchmark their economies with competitors globally, and there are many ways to quicken the development of profitability. Banks and financial service firms and related

administration organizations in the developing countries should adopt online payment systems to get e-trade and value speculation. Tourism and its Internet incarnation are routinely referred to as some of the quickest developing e-commerce sectors [73].

Figure 2.6 Shown, online visiting browsing is the maximum rates in Latin America, but online buying is the minimum rates according to [74]. Also, online buying products in Asia-Pacific are the maximum of both of them [74]. In general, we can say the size and growing of e-commerce in Asia-Pacific are better than the others.

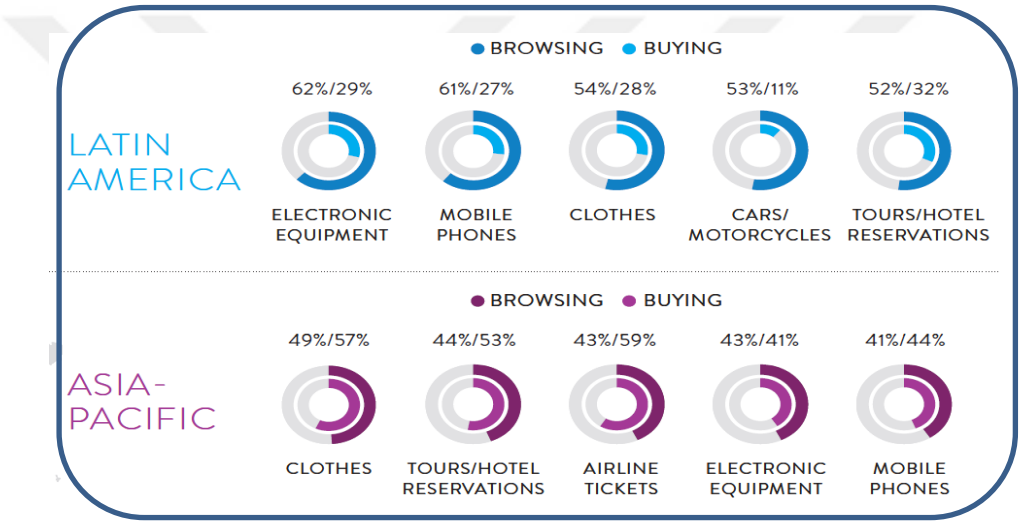

Figure 2.6. Browsing/Buying Categories [74]

Table 2.2 includes nineteen products which are ordered from top to bottom. The table demonstrates the rate of growing these nineteen products from 2011 to 2014; it means that between these two years the numbers are changed. The rate of growing products was recorded from the highest to the lowest. The rate of growing event tickets in 2011 was 22%, but in 2014 the rate was developed to 41%, it means 19% is the amount rate which was increased between these two years [74].

Table 2.2. Growth Rate Online Shopping 2011-2014 [74]

	2011	2014	PERCENTAGE POINT GROWTH
EVENT TICKETS	22%	41%	19%
E-BOOKS	15%	34%	19%
COMPUTER SOFTWARE	9%	27%	18%
TOURS AND HOTEL RESERVATIONS	27%	44%	17%
SPORTING GOODS	15%	31%	16%
TOYS AND DOLLS	13%	29%	16%
AIRLINE TICKETS AND RESERVATIONS	33%	48%	15%
VIDEOS, DVDS AND GAMES	15%	28%	13%
MUSIC (NOT DOWNLOADED)	14%	27%	13%
COMPUTER HARDWARE	18%	30%	12%
BABY SUPPLIES	8%	20%	12%
FLOWERS	7%	18%	11%
CAR, MOTORCYCLE AND ACCESSORIES*	8%	17%	9%
ALCOHOLIC DRINKS	8%	17%	9%
ELECTRONIC EQUIPMENT	27%	34%	7%
COSMETICS*	25%	31%	6%
HARDCOPY BOOKS	34%	39%	5%
GROCERIES	22%	27%	5%
CLOTHING, ACCESSORIES AND SHOES	42%	46%	4%

The percentages change from the visitors to others. Table 2.3 shown, includes five types of generations, to browse and buy online shopping, age 21-34 are recorded the highest level and generation (65+) are the lowest level [74].

Table 2.3. The Differences Visitors' Percentage [74]

Visitors	Browse Online Shopping	Buy Online Shopping
Generation (Under 20)	6%–9%	5%–9%
Millennial (21 – 34)	49%–59%	52%–63%
Generation X (35 – 49)	25%–28%	25%–30%
Baby Boomers (50 – 64)	7%–13%	6%–13%
Silent Generation (65+)	1%–3%	1%–3%

Table 2.4 displays that the size of trade e-commerce has advanced in the date. Also clearly China has the highest scores among the top 30 countries in the world, when, compared with the other countries this table includes the top 30 countries from China to Malaysia [75].

Table 2.4 Countries level E-Commerce [75]

Rank	Country	Online Market Size 40%	Consumer Behavior 20%	Growth Potential 20%	Infrastructure 20%	Online Market Attractiveness Score
1	China	100,0	68,8	100,0	51,1	84,0
2	Japan	100,0	100,0	17,4	99,1	83,3
3	United States	100,0	77,6	39,8	96,5	82,8
4	United Kingdom	100,0	77,5	14,7	86,3	75,7
5	South Korea	79,6	97,4	93	951	72,2
6	Germany	90,3	78,3	28,1	65,1	70,4
7	France	85,5	75,7	7,4	71,6	65,2
8	Brazil	37,2	51,2	64,7	64,1	50,9
9	Australia	15,7	89,4	46,2	86,9	50,8
10	Canada	17,7	73,5	48,3	915	49,7
11	Singapore	2,3	93,1	28,9	100,0	45,3
12	Argentina	9,2	59,1	75,7	68,0	44,2
13	Russia	34,9	51,8	56,4	42,3	44,1
14	Hong Kong	3,2	93,7	17,2	100,0	43,4
15	Italy	16,1	52,2	64,3	60,7	41,9
16	Sweden	12,1	77,5	21,7	85,7	41,8
17	Slovakia	2,0	71,5	86,4	44,3	41,2
18	New Zealand	2,5	92,3	28,1	78,5	40,8

19	Netherlands	16,2	77,5	17,4	73,9	40,2
20	Chile	3,9	61,0	56,5	74,8	40,0
21	Finland	13,3	77,2	13,6	82,1	39,9
22	Turkey	10,7	26,6	72,9	78,4	39,9
23	Venezuela	2,5	49,5	100,0	42,1	39,3
24	Belgium	9,8	70,6	26,5	73,1	38,0
25	U.A.E.	0,9	50,3	49-2	87,8	37,8
26	Norway	12,3	77,5	97	75,7	37,5
27	Ireland	7,2	62,3	42,2	67,9	37,4
28	Denmark	10,2	78,3	14,1	73,0	37,2
29	Switzerland	13-2	68,2	10,9	79,4	37,0
30	Malaysia	1,0	63,0	44,2	75,0	36,8

3. PRODUCT SEARCHING FOR E-COMMERCE AND METHODS

With the quick development of e-commerce, services for online searches for products have appeared as a common and active model for customers to select a transaction after looking for their favorite products. Nowadays, search engines of products depend on the relevant model's adaptation, which is devised for information retrieval. There is still an enormous difference between finding the most wanted products and retrieving highly recommended products [76].

3.1. Text Search

A text search engine is a type of search that helps customers or users find products by retrieving information from the e-commerce website or the Internet. The full-text search is a popular searching method that depends on metadata or on the original text parts signified in databases. A full-text search is used for comparing each word of the search request with each word inside the database. This kind of search is everywhere on the Internet and contains the kind of common search for language we normally look for in e-commerce search engines. It requires a text as input to search for products [77]. In the following paragraphs, we explain the disadvantages of the text search:

The synonym problem may be the greatest and most popular weak point of text searching. This issue occurs because there is more than one way to look for a name or express a certain idea. There are many aspects of the disadvantages of synonyms. Different spellings of words that have similar meaning can sometimes be spelled in various forms, such as the many American and British spelling deviations. It means that retrieving products has less accuracy with different spellings than with exact synonyms [77]. The short version of terms, such as abbreviations, can present issues in the text search method because a product may include the short or long form. While searching for products, different languages might be used, which the customer is unable to match, and this includes the version of the foreign language for that concept, unless the two terms happen to be cognates [77].

3.2. Image Search

This is a new form of content retrieval built with the help of an image search engine. One can also look up similar images based on the source image offered to the search engine. It is defined as a search engine designed to get information based on the input of an image [78].

The image search is mostly used by e-commerce buyers and sellers and to search for more information on the image of an unknown item. The image search has developed into a common feature in many search engines. This idea is to show the right products on e-commerce for online shopping to customers. Moreover, they are some important advantages to image searches.

Image search has the capability to extract details of good marks on a wide scale with accuracy. It is in every shop's best interests to confirm all details about the product clearly. The product catalogue helps find products both on the website and in search engines. In addition, it helps customers find products and decide which product they want because they see it before buying.

It is very easy and could save a lot of time by finding correct results. A search of image choices makes it important for buyers to combine search choices to get the best search results faster with the facility of introducing more new faces in the future.

Since E-Commerce includes one of the quickest developing sections on the Internet, and online search products have newly developed a viable way for searching customers' chosen product. Looks like to common purpose Web search, product search engines permit customers to submit keyword-based queries and return to them with a list of products, in which customers can choice the products they want to buy online.

The most of the product search engines today are made based on the relevance model from classic information. In chapter four, we have more explained about it [76]. Many papers and researchers already have been reviewed and the limits of the earlier methods were introduced:

3.2.1. Wavelet-Based Image Indexing and Searching

Wavelet-based image indexing and searching describes indexing a new image and using a retrieval algorithm with the ability to search for a partial image in a large image database. The aim of this algorithm is to characterize the color variations over the spatial extent of the image in a manner that supports semantically meaningful image comparison [79].

Preprocessing the Images in the Database

Variant color formats of an image are currently has been used and the most widely used formats. The researchers have been the first time normalizes the data. Database testing of relatively small images, rescaled thumbnail consisting of 128*128 pixels in RGB color space is adequate for the purpose of computing the feature vectors [79].

Considering that color distances in RGB color space are not reflecting the actual human perceptual color distance, they have been converted and stored the image in a component color space with intensity and perceived contrasts, the authors have been defined the new values at a color pixel based on RGB values of an original pixel as follows:

$$\begin{cases} C1 = (R + G + B)/3 \\ C2 = (R + \ max - B))/2 \\ C3 = (R + 2 * (max - G) \ + B)/4 \end{cases} \quad (3.1)$$

$$\begin{cases} RG=R-2*G+B \\ BY=R-G+2*B \\ WB=R+G+B \end{cases} \quad (3.2)$$

Where (*C1, C2, C3*): photometric invariant color features. (*R, G, B*): sensor color space. (*RG, BY, WB*): Three opponent color axes. Here, max means the value of maximum possible for a component of each color in the RGB color space. For a 24-bit standard

color image, max=255. Apparently, every component of the color in the range of the new color space from 0 to 255 as well. This color space is same as the axes of an opponent color [79].

3.2.2. Image Similarity Search Using NSA

The Negative Selection Algorithm (NSA) is an immune-inspired algorithm that could be used for many purposes, such as fault detection, data integrity protection, and virus detection. The original NSA was inspired by the way natural immune systems distinguish the self from the others [80].

Direct Pixel Similarity (DPS)

The earliest idea researchers explored is direct pixel similarity (DPS), which works at the pixel level. Thus, the simplest and most apparent similarity measure is a direct red, green, and blue (RGB) comparison of each pixel. If the pixels are same, they match details. All the matched cases are counted and normalized to one. As we can see, it is not a good detector, as it is unusual that many pixels in a picture match perfectly with the target picture [80].

There might be changes in brightness or other minor differences that are barely visible and can be noticed which negatively influence such similarity measures. Thus, the similarity match by considering a range of each pixel value v, such that we still agree on a match if the value v of each color component is in the range $[v - r, v + r]$, for a given parameter r.

Also, they have defined a match on a group of n by n pixels if, for a certain threshold t, t pixels in a group of the target image match the foreign image. At the end, we count completely matched and divided pixels by the total amount of pixel groups that were compared to find the final measures of similarity.

3.2.3. Comparative Analysis of Image Search Algorithm

Content-based image retrieval defines a process to find the exact similar picture(s) in the database of the image since a query image is given. In CBIR, the retrieval of an image is based on the similarities in their content.

Due to the increasing demand to search for the digital images efficiently and accurately, existing image search engines on the web, for example, Google and Yahoo, are both based only on relevant text. However, some relevant texts lead to serious limitations. To solving these problems, the researchers have used QBIR [81].

In this research, the researchers have used four techniques: the average of RGB, local color histogram (LCH), global color histogram (GCH), and color moment of hue, saturation, and brightness value (HSV). The aim of the authors in using these four techniques is to evaluate and choose the best one [81].

A) Average RGB

Average RGB is for computing the average values in R, G, B channels of every pixel in an image and used as the descriptor of an image for comparison purpose. The following four equations are for computing the average *R, G, B* component of an image *I*:

$$r = \sum_{x=1,y=1}^{x=w,y=h} \frac{r(I(x,y))}{w*h} \tag{3.3}$$

$$g = \sum_{x=1,y=1}^{x=w,y=h} \frac{g(I(x,y))}{w*h} \tag{3.4}$$

$$b = \sum_{x=1,y=1}^{x=w,y=h} \frac{b(I(x,y))}{w*h} \tag{3.5}$$

$$d(I_a, I_b) = \frac{\sqrt{(ra-rb)^2+(ga-gb)^2+(b-bb)^2}}{3} \tag{3.6}$$

Here the distance of an image I_a and I_b, where the weighted Euclidean is used. The measured distance between two exact images will be 0 and the distance between to most non-similar images (black and white) will be 1 according to the range of RGB from 0-1 or 0-255 [81].

Notation:

I: an image.

W: width of the image *I*.

h: height of image *I*.

I(x,y): the pixel of the image *I* at row *Y*, column *X*.

R (p), G (p), B(p): the red, green and blue color component of pixel *p*.

r_a,g_a.b_a:the average red, green and blue component of image I_a.

$d(I_a, I_b)$:the distance measure between image I_a *and* I_b [81].

B) Local Color Histogram (LCH)

Usually, images are represented in RGB color space and some of the most important bits are used for each color channel. Color Histogram is one of the widely used techniques for color feature extraction in color-based image retrieval. Color Histogram is a method for describing the color content of the image, constructed by counting the number of pixels of each color [81].

L_1-distance

$$|H_I - H_I'| = \sum_{j=1}^{n} |H_I[j] - H_I'[j]| \quad (3.7)$$

L2-distance

$$||H_I - H_I'|| = \sum_{j=1}^{n} (H_I[j] - H_I'[j])^2 \quad (3.8)$$

Default vector comparing method: by default, measuring the distance of 2 images as follow:

$$H_I - H_I' = \sqrt{\frac{\sum_{j=1}^{n} (H_I[j] - H_I'[j])^2}{n}} \quad (3.9)$$

The color histogram H_I is a compact summary of the image. A database of images can be queried to find the most similar image to *I* and can return the image *I'* with the most similar

color histogram H_I. Typically color histograms are compared using the sum of squared differences (L2-distance) or the sum of absolute value of differences (L1-distance). So the most similar image to I would be the image I' minimizing the L2-distance or L1-distance. Note that we are assuming that weighted evenly across different color buckets for simplicity [81].

Notation**:**

$H\ (h_1, h_2, ..., h_n)$: a vector, in which each component hj is the number of pixels of color j in the image.

n: number of distinct(discretized) color.

I: an image.

H_I: the color histogram of the image I [81].

C) Global Color Histogram (GCH)

This technique is represented images with a single histogram and it does not capture the content of images adequately. Histogram Pattern a color histogram represents the distribution of colors in an image, through a set of bins, where each histogram bin agrees to a color in the quantized color space using 1 vector, $H\ (h_1, h_2, ..., h_n)$ to describe an image. As shown in Figure 3.1.

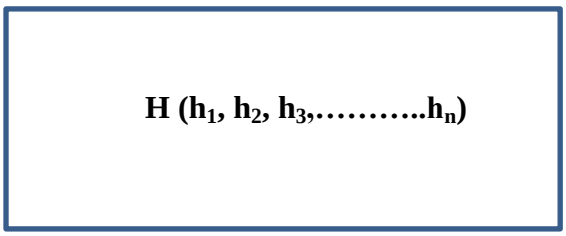

Figure 3.1. GCH Image Describing Vector [81]

$$h_j = \frac{\text{Number of pixels of color } j \text{ in the image}}{\text{Total number of pixels in the image}} \quad (3.10)$$

Where i is the color bin in the color histogram and $H[i]$ represents the number of pixels of color i in the image, and n is the total number of bins used in color histogram [81].

D) Color Moment HSV

The basis of color moment lays in the supposition that the distribution of color in an image can be interpreted as a probability distribution. Probability distributions are characterized by a number of unique moments, e.g. normal distributions are differentiated by their mean and variance. It, therefore, follows that if the color in an image follows a certain probability distribution, the moments of that distribution can then be used as features to identify that image based on color [81]. The three color moments can be defined as:

Moment 1– Mean: Mean can be understood as the average color value in the image.

$$E_j = \sum_N^{j=1} \frac{1}{N} p_{ij} \tag{3.11}$$

Moment 2 – Standard Deviation:

$$\sigma_i = \sqrt{\frac{1}{N}\sum_N^{j=1}(P_{ij} - E_i)^2} \tag{3.12}$$

The standard deviation is the square root of the variance of the distribution.

Moment 3 – Skewness:

$$S_i = \sqrt[2]{\frac{1}{N}\sum_N^{j=1}(P_{ij} - E_i)^3} \tag{3.13}$$

Skewness can be understood as a measure of the degree of asymmetry in the distribution. The Similarity function between two image distributions is defined as the sum of the weighted differences between the moments of the two distributions [81]. Formally this is:

$$d_{mom}(\mathrm{H},\mathrm{I}) = \sum_{i=1}^{r} W_{i1}|E_{i^1} - E_{i^2}| + W_{i2}|Q_{i^1} - Q_{i^2}| + W_{i3}|S_{i^1} - S_{i^2}| \tag{3.14}$$

Pairs of images can be ranked based on their d_{mom} values. Those with greater values are ranked lower and considered less similar than those with a higher rank and lower d_{mom} values.

Notation:

(H, I): Are the two image distributions being compared.

i: Is the current channel index.

r: Is the number of channels.

E_{i^1}, E_{i^2} are the first moments (Mean) of the two image distributions.

Q_{i^1}, Q_{i^2} Are the second moments (Standard Deviation) of the two image distribution

S_{i^1}, S_{i^2} Are the third Moments (Skew-ness) of the image distributions.

W_i : Are the Weights for each moment [81].

3.2.4. Image Search Engine Using SIFT Algorithm

The approach that the scale-invariant feature transform (SIFT) feature detection takes in this implementation is similar to that taken by the researchers, which is used for object recognition. By this work, the invariant features extracted from images might be used to perform dependable matching between various views of an object [81].

The aim of this implementation focuses on all features from an image and on trying to use these features to perform the image search. The researchers also used RGB and RGB details for image retrieval for implementation of the study. We focus the work on this SIFT algorithm. In the RGB color model, a color image can be represented by the intensity function:

$$I\,RGB = (FR, FG, FB), \tag{3.15}$$

Where *FR(x, y)* is the intensity of the pixel *(x, y)* in the red color, *FG(x, y)* is the intensity of the pixel *(x, y)* in the green color, and *FB(x, y)* is the intensity of the pixel *(x, y)* in the blue color. The intensity of each color is stored using 8 bits [82].

3.2.5. Content-Based Image Retrieval System Based on PRESS

In this work, the researcher's shape features have been extracted from the database images and the same are polar and raster scanned into specified intervals in both radius and angle, using the proposed polar-raster edge sampling signature (PRESS) algorithm. The aim of the technique is to find imperfect instances of objects within a certain class of shapes using a voting procedure [83].

3.2.6. Content-Based Image Retrieval Using Color and Edge Direction Features

In this paper, the researchers present a novel technique that employs two of the edge and color direction features for CBIR. In addition, the images are first divided into sub-blocks of the same size, and later, the edge and color direction features of each sub-block can be extracted. Then, a codebook of the color features is constructed using a clustering algorithm. Next, each sub-block is mapped to the codebook.

Finally, color index codes are used to retrieve images, and the edge direction feature is used as the color feature weight, which belongs to the same color feature sub-block [84].

The researchers use the key block-based image retrieval algorithm, which is a generalization of the text-based information retrieval methods.

Similarity Measure

After getting the colors features of direction, they have calculated the distance between search images and the query image**.** And they presented a key block-based of similarity measure which joins the edge direction feature into the feature of color similarity measure function [84]. The similarity measure defined in the following formula:

$$S(q, q_n) = \frac{1}{1+dis(q,q_n)} \tag{3.16}$$

$$dis(q, q_n) = \sum_{i=1}^{t} \frac{S(q,q_n)_i * |W_{id} - W_{iq}|}{1+W_{id}+W_{iq}} \tag{3.17}$$

$$s(q, q_n) = \frac{|f_{0d} - f_{0q}| + \left|f_{1d} - f_{1q} + \cdots |f_{5d} - f_{5q}|\right|}{\max(f_{0d} \ldots f_{5d}, f_{0q} \ldots f_{5q}} \tag{3.18}$$

Where q is the query image, $D=\{d_n\}$ is the search image data,W_{id} is the frequency that the color code K_i appear in image d_n, W_{iq} is the frequency that the color code K_i appear in image q ,and | W_{id} - W_{iq}| is the comparison of the frequency which sub-block has the same color code , is the distance function that compares the edge direction which belongs to the same color code[84].

4. IMAGE RETRIEVAL AND APPLICATION

Due to the growth of online shopping customers on the Internet, several collections of digital images have been developed. In this part, we focus on some important details, including the information about the RGB color model and how to use RGB color to retrieve images for online shopping. Moreover, we describe the relation of RGB colors and how to find similar images in the database, and we compare images between query images with the images in the database. In this part, we present the general architecture of our thesis, the implementation, the mechanisms, and which program languages were used.

4.1. Image Retrieval

Retrieval of the image is done by a computer system for retrieving images, browsing, and searching a huge database of digital images. The most popular and traditional methods for image retrieval use some methods to describe the images so that retrieval might be performed using annotation words. In this thesis, we have used RGB, which is a good representation option. Nowadays, RGB is usually used for many properties [85].

4.1.1. RGB Method

The RGB color is a collective color model that includes red**,** green, and blue light mixed together in several ways to produce a wide display of colors. The essential aim of the RGB color model is for sensing, representing, and showing images in electronic systems, like e-commerce. However, it has been used in many other things [81].

4.2. How to Work RGB Color to Retrieval Image

The system is made for comparative image recovery in online businesses. The online shopping system is used for enhancing system client understanding. We have designed a system the user can understand and use easily for a broad e-commerce application, such as image search and item data-gathering subsystems.

This system may give a direct contact between images in the database and our online shopping, on which purchasers can automatically and directly search the same images as indicated by the images from the data stage. Simultaneously, it can be utilized to give exactness of online advertising to companies. In Figure 4.1 the proposed system component steps are explained:

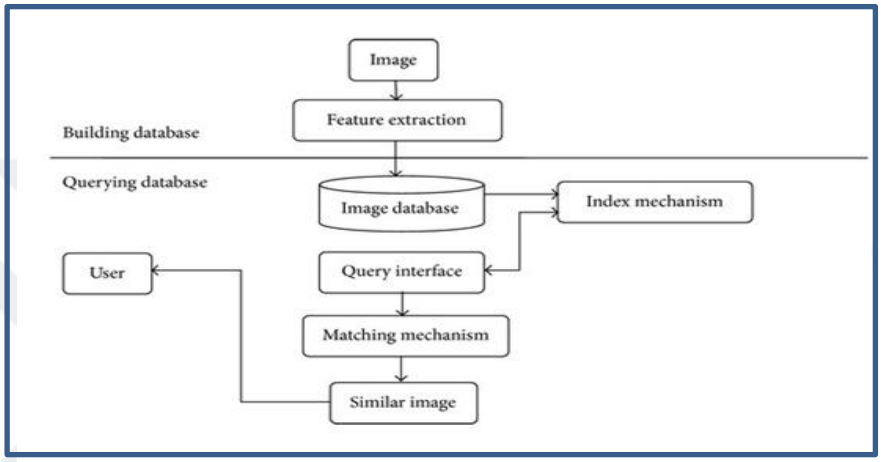

Figure 4.1. The Basic Components of the Similar Image Retrieval [86]

4.2.1. Feature Extraction

Feature extraction is a process for making the representation from the original. Commonly, feature extraction is used to find different image segments. Image features include basic features and semantic features. The basic features of the image, which are perceptible, include the color. The semantic features of the image, which describe the image content, are qualitative. They are extracted artificially or by human-computer interaction. Features of images are classified as follows:

Color

Color information distribution can be derived from RGB. The RGB projection determines the similarity phase of the studied work where it compares the query image with that of the images in the feature database. Moreover, RGB projection is a technique that appraises the image vertically and horizontally. The arrangement is built on the similarities between the target and the non-target images.

The detailed steps are explained below:

Step 1: Test image is the projection onto each target and non-target image.

Step 2: Makes relation RGB color.

Step 3: The target images are compared with the sum of the images belonging to the non-target images to determine similar images.

Step4: The compared image is used to identify similar images automatically.

Step5: It is tested against the whole collection of the database.

Step6: Image pairs along with their similarity percentages are returned.

4.2.2. Color Histogram

It represents the division of colors in an image, where every histogram bin matches a color in the quantized color space. A color histogram is a set of bins where any bin represents an appropriate color of the color space being used. The number of bins depends on the number of colors in the image. A color histogram for a given image is well defined as a vector [87], as shown in Equation (4.1):

$$H = \{ H[0], H[1], H[2], H[3] \ldots\ldots\ldots H[i], \ldots\ldots H[n] \} \quad (4.1)$$

Where *i* represent the color bin from the color histogram and *H[i]* represents the number of pixels of color *I* from the image, and *n* is the total number of bins used in the color histogram.

In addition, in Equation (4.2), the histogram intersection method is used to measure the distance *S* between the query image *Q* and the image *P* in the image database [87]:

$$s = 1 - \sum_{i=0}^{255} min(H_Q[i], H_P[i]) \quad (4.2)$$

4.2.3. RGB Color Relation Index

In this thesis, the relation *r* is a relation between the query image and retrieving images. When *A* and *B* are the query image and retrieved image, respectively, they are reduced to the same size. The relation *r* is defined in the following [88]:

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)\ (\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \quad (4.3)$$

Where $\bar{A}$ is the average of element *A* and $\bar{B}$ is the average of element *B*. Then, the images have been indexed in a database based on these features. According to the sample image, the user can retrieve images in the database using some similarity measurements, and the system provides a similar outcome to users. In addition, this mechanism is intermediate between the query image and images already stored in the database. It is provided through the user interface (UI) containing query by keyword, browsing categories, etc. [88].

Let *A* = query image and *B* = retrieved image:

$$A = \begin{pmatrix} R=f(xi\ ,\ yi) \\ G=f(xi\ ,\ yi) \\ B=f(xi\ ,\ yi) \end{pmatrix} \quad (4.4)$$

BP = resize (retrieved image *P* order); size (256*256)

$$BP = \begin{pmatrix} R'=f(xi\ ,\ yi) \\ G'=f(xi\ ,\ yi) \\ B'=f(xi\ ,\ yi) \end{pmatrix} \quad (4.5)$$

Where *(xi, yi)* is the *i*th pixel of the classified image for (*p* = 1, …, *n*), where *n* is the number of pictures. In their arrangement, all pixels in the classified image were represented by a vector in three primary-color spaces: red (R), green (G), and blue (B).

In this study, we have used RGB color components and relation statistics to index the similarity image for image retrieval, using these formulas [88]:

For (q=1 to n(number of pixel=256*256)

$$\text{Relation (r)} \quad A,B\,P=\begin{pmatrix} R=f(xi\,,\,yi) \\ G=f(xi\,,\,yi) \\ B=f(xi\,,\,yi) \end{pmatrix}, \begin{pmatrix} R'=f(xi\,,\,yi) \\ G'=f(xi\,,\,yi) \\ B'=f(xi\,,\,yi) \end{pmatrix} = \begin{pmatrix} r_R' \\ r_G' \\ r_B' \end{pmatrix} \tag{4.6}$$

$$r_{\mathrm{R}} = \sum_{q=1}^{q=n} r_{\mathrm{Rq}} \tag{4.7}$$

$$r_{\mathrm{G}} = \sum_{q=1}^{q=n} r_{\mathrm{Gq}} \tag{4.8}$$

$$r_{\mathrm{B}} = \sum_{q=1}^{q=n} r_{\mathrm{Bq}} \tag{4.9}$$

Where *Rindex* is an RGB color relation index for the query and retrieved image. The value of this index is an indication of the degree of the relation, and depending on application, it can be used to make a proper retrieval decision. The images are ordered by the RGB color relation index to show the corresponding images [88].

Rindex is the vector summation of RGB relation described by this equation:

$$Rindex = \sqrt{r_R{}^2 + r_G{}^2 + r_B{}^2} \tag{4.10}$$

4.2.4. Similarity Calculation

Image retrieval calculates similarities between two images. Thus, the outcome for retrieval is not a single image; it generates a list of images that are ranked according to the form of similarities. For similarity calculation for retrieval of the image, the following formula (Euclidean distance) [89] is used:

$$D(A,B) = \sqrt{\sum_{i=1}^{d}(A_i - B_i)^2} \tag{4.11}$$

where A is the query image, B is the target image, and i is the pixel value. After the image has been uploaded, the similarity distance will be checked and compared with the already-stored images inside the database. Later, we obtain the value for each uploaded image and search for the similarity between the value and the stored image, and if the condition is approved, then we find the image.

Retrieval of a similar image, which is actively done by the user, can support the user to find the products faster and Structural Similarity SSIM another type is used to find similarity between both of image query with images in database index is a novel method for measuring the similarity between two images. The SSIM index can be viewed as a quality measure of one of the images being compared provided the other image is regarded as of perfect quality [90]. Similarity calculated codes:

```
public double CalculateSimilarity(RgbProjections compare)
        {
            var hS = cps(horizontalProjection,
compare.horizontalProjection);
            var vS = cps(verticalProjection, compare.verticalProjection);
            return Math.Max(hS, vS); }
private static double cps(double[] source, double[] compare)
        {
            if (source.Length != compare.Length)
            {
                throw new ArgumentException();}
            var frequencies = new Dictionary<double, int>();
for (var i = 0; i < source.Length; i++)
            {
                var difference = source[i] - compare[i];
                difference = Math.Round(difference, 2);
```

```
                difference = Math.Abs(difference);
                if (frequencies.ContainsKey(difference))
                {
                    frequencies[difference] = frequencies[difference] + 1;
                }
                else
                {
                    frequencies.Add(difference, 1);    }            }
            var   deviation   =   frequencies.Sum(value   =>   (value.Key   *
value.Value));
       deviation /= source.Length;
            deviation = (0.5 - deviation) * 2;
            return deviation;    }}
```

The objective of this query is to powerfully search for and retrieve images from the database that are the same or similar to the query image contributed by the buyer. Resulting in good measures of similarity between images based on various feature sets is a testing work. Measurement of similarity is the operation of assessing the difference/similarity between the query image and the database images using their features.

The system coding consists of four classes for each search image inside the database to compare RGB. If the image that has been searched for existed in the database or had the most similarities, for example, 90% or more, then the result will be shown at the top of the search results. For this thesis, we have used two process image parameters, which are called (fileinfo[ ]files, fileinfo target):

- FileInfo[] files: is the list of all the images that have been sent to the database.
- FileInfo target: is the image that is compared with the list of the images in the database (i.e., FileInfo [] files). Thus, it will retrieve all the images in the database, which resembles 90% of the image that has been used in the search engine. This range is not fixed; we can change it according to our work from the high to low range.

There is a class called ComparableImage to calculate the match by passing each of them through the calculation. Compare image codes:

```
var similarity = source.CalculateSimilarity(destination);
var sim = new SimilarityImages(source, destination, similarity);
if (sim.Similarity >= 90)
{
ListImageNames.Add(destination.File.Name);
}
```

4.3. Mechanism

In this thesis, the implementation is based on some programming languages and techniques, as follows: ASP.NET model-view-controller (MVC) is used for the web application structure that is supported by Microsoft, which executes the MVC design. The programming depends on ASP.NET, and ASP.NET MVC permits software designers to fabricate a web application as a composition of three parts: model, view, and controller. The MVC model characterizes web applications with three rationale layers:

- Model (business layer)
- View (display layer)
- Controller (input control)

How to Use the System

For system integration the system is using the MVC framework model, extending then processing data through view point, data will be read it in the viewpoint by the controller later the MVC will send the model to input data of user's control, the model object is responsible for accessing the database. Using the controller decoupling performance logic and business logic, therefore to improve system maintainability, reusability, and scalability, effective separation of business logic, interface display implementations and data [91]. Figure 4.2 displays, that how MVC works together. The MS SQL Server is used to implement a database in this thesis because it is secure and can process many users at the same time.

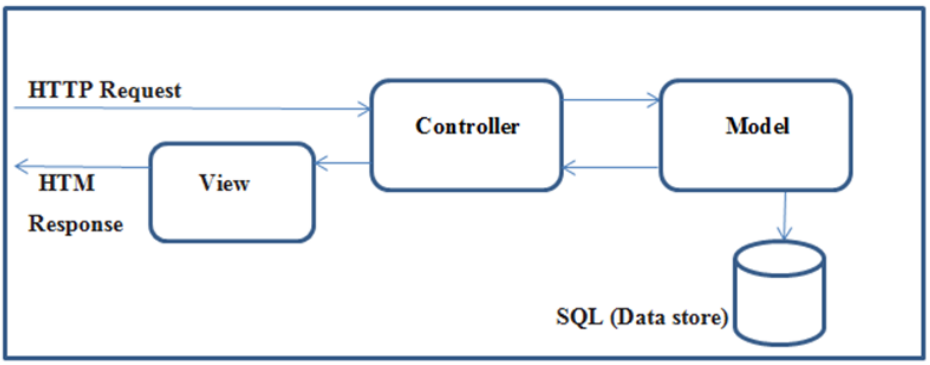

Figure 4.2. MVC Structure Query [91]

In this thesis, MVC application that stores images in MS SQL Server 2014 has been used. The user retrieves the images and displays them in an MVC view. Based on my Internet research, this has been discussed many, many times, and there are lots of approaches to do this. In this thesis, is stored the images in the database. As I said, there are many approaches to retrieving the image and displaying it in a view. In the following, have mentioned general code about MVC:

View Model:

```
public class ImageViewModel
 {
     public byte[] Image { get; set; }
     public MemoryStream Imagestream { get; set; }  }
```

Controller:

```
public ActionResult Index()
{
    DataAccess.DataClasses1DataContext ctx  = new
DataAccess.DataClasses1DataContext();
    using (ctx)
    {
        var db  = (from q in ctx.tblImages
                   select  new Models.ImageViewModel
                   {
                       Image = q.ImageData.ToArray()
                   }).ToList().Take(5);
        List< Models.ImageViewModel > ImagesList= new List
<Models.ImageViewModel>();
        foreach (var item in db)
        {
            Var stream = new MemoryStream (item.image);
            Models.ImageViewModel obj = new Models.ImageViewModel();
            obj.Imagestream = stream;
            ImagesList.Add(obj); }
           return View(ImagesList);
}               }
```

View:

```
@model IEnumerable<MvcApplication3.Models.ImageViewModel>

   @foreach (var item in Model)
   {
       <img src ="@(item.Imagestream).png" />   }              }
```

Search image in this system will depend on ASP.Net framework and also it can be used on visual programs either, but because this work has been created for images, that is why visual C#.Net; it has more compatibility for such things, because of security and flexibility.

4.4. Project Architecture

Our system has three-layer architecture to make the application more understandable and easy to maintain and modify, for the application that requires some modifications and implementations for different clients like for the shopping sites. In the following, has been explained for each layer:

Presentation layer

The first layer (presentation layer), the ASP.NET webforms, is the top layer, including the web server, and our resources (directories, files, images, and another type of information), which all are located on the web server. In addition, it is the best level of the application in the UI. It shows information that can be used for facilities like shopping, browsing merchandise, and purchasing. It includes the user UI components of the site and contains all the logic that manages the communication between the buyers and visitors.

The important tasks of the UI are to translate results to help the buyer better understand. It will be opened to the end user as a usual web application, which contains ASP.NET web forms, styles (CSS), and scripts (client side JavaScript).

Logic Layer

The second layer is the logic layer; it has been created for our thesis. It manages process commands and the applications, makes logical decisions and estimations, and makes calculations. This layer also moves and processes data between the two other layers. The logic layer controls the application functionality by executing complete processing. It receives an offer from the first layer and returns a result to the same layer depending on the logic layer it includes.

Database layer

The third layer is the database layer, where all the data, such as the scheduling table, profiles, images, and other detailed information are stored in the database using the SQL server. This information is stored and retrieved from the database or file system. The information is later processed back through the logical layer and finally back to the user database for storing the data in the form of tables with the ability to establish relationships among them. Figure 4.3 explains from the first layer to the end layer.

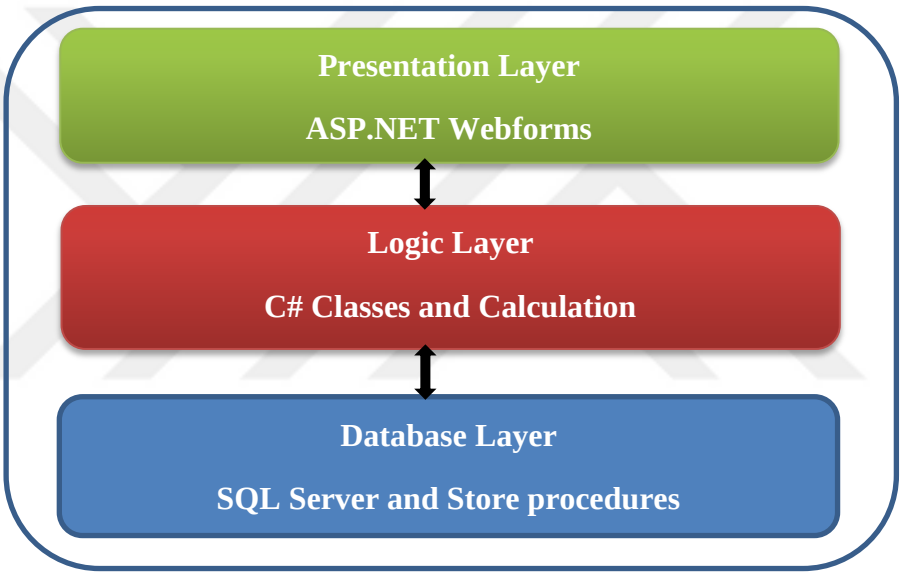

Figure 4.3. Project Architecture [92]

4.5. Visual Studio 2013

In this thesis, we used Microsoft Visual Studio 2013, which clears the graphical user interface (GUI) organization and 'reduces clutter and complexity'. It has integrated support for developing Microsoft applications and enhancing testing tools and design, including an interactive designer. It supports various programming languages and permits the debugger and code editor to support nearly any program.

Properties or Features of Visual Studio

- Code Editor enhancements, VS code editor, supports setting bookmarks in code for speed navigation
- Better Debugging and Performance.
- Rapidly create properties
- Server Explorer used to administrator database connections on an available computer and used to browse running Windows Services**.**
- Designer, Visual Studio contains a host of visual designers to help in the growth of applications.

4.6. Programming Language and GUI

The GUI is a visual computer environment that represents programs, which is how the user interacts with the database through programmed components. The GUI for the current study is designed using Microsoft Visual Studio 2013 (IDE) Web application with visual C#. Through this program, both Windows form applications and web applications can be developed. For our thesis, we used some powerful programs, which have been shown and explained in Figure 4.4.

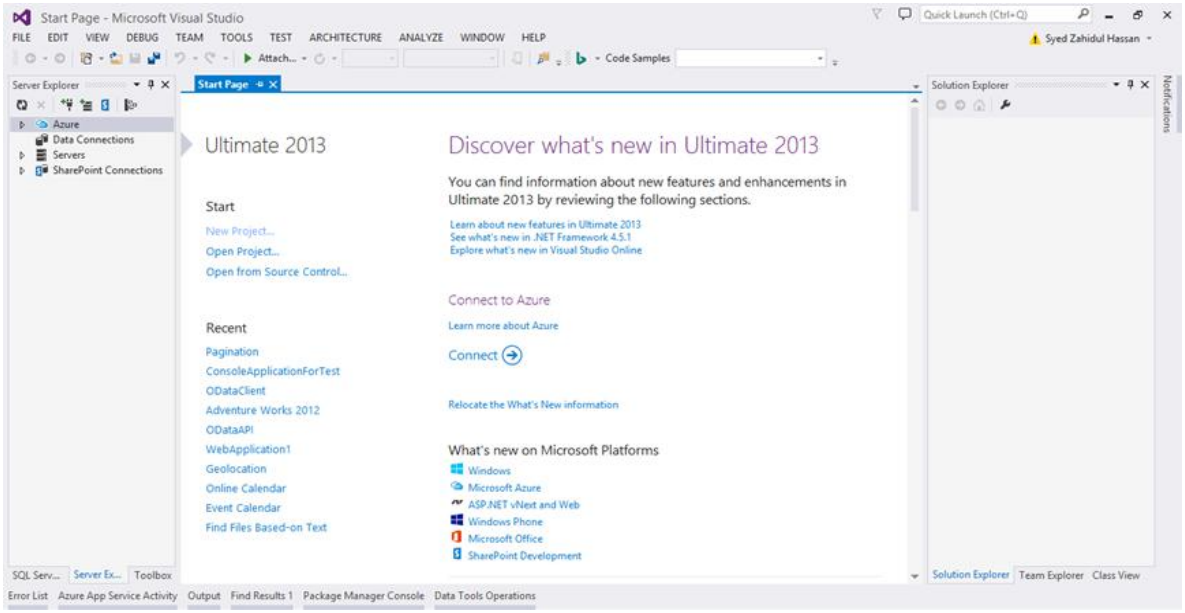

Figure 4.4. Microsoft Visual Studio Start Page [93]

C#

The language C# (see-sharp) is a multi-purposed computer programming language convenient for all development requirements and security. Although C# is derived from the C programming language, the Sharp language is directed towards programmers and C programmers as an extension of these languages with more speed than in C or C++. Moreover, C# is intended to be a modern, simple, object-oriented, general-purpose programming language. The rich class library makes many functions easy to implement, providing good support for the distributed system. Usually, it is much more efficient and runs faster. It can provide a great amount of flexibility in the designs [94].

Web Application

In this thesis, we used ASP.net by C#; ASP.NET is a server-side web application framework designed for web development to produce dynamic webpages. It allows programmers to build web services, dynamic websites, and web applications. It was developed by Microsoft. We chose ASP.NET because it has many advantages, such as:

- Nearly all processes are managed and controlled by ASP.NET runtime, if a process is dead, a new process may be created in its place, which protects the application and makes it continuously accessible to handle requests.
- The ASP.NET framework supports varied languages.
- ASP.NET provides the best performance by taking improvement of early binding, just one-time collection, and native optimization.
- Large and improve controls features and class library.
- Optimum speed of growth.
- It is developed applications are available on a global basis leading to active information management
- Strict security.
- Can easily configuration of the application.
- High accuracy.
- ASP.NET drastically decreases the amount of code need to create large applications.

4.7. Database

The first step in the development of our thesis was to design a robust database. It is the basis for all data that are stored in the form of tables. In this thesis, the database comprises many tables. Each one is used for a special purpose for storing data. One of the most important advantages of this database is sharing the data among several applications at the same time.

There are many types of database management systems. For this thesis, we used MS SQL Server 2014 [95].

Microsoft SQL server is a strong relational database engine, which is used to store and maintain data in distinct levels. The SQL server can query many rows from one or more tables using just one statement; this feature is one of the most important advantages of SQL.

In the thesis, we used a database because one can store data and extract it whenever required. In addition, one can store and extract data from a database in various ways. There are some advantages of a database system [95]:

- An organized and comprehensiveness of recording the result of the firm's activities.
- Data integrity and independence from an applications program.
- Enhanced data access to users through the use of host and query languages.
- Enhanced data security.
- Easy development of new applications program.
- Speed

Advantages of MS SQL Server

Because SQL Server consists of many features in the following we have explained:

- Ease of use and maintenance.
- Fast and efficient.

- Security and stability.
- Availability and Reliability.
- Accuracy
- Facilitated development of new applications program.

5. IMPLEMENTATION E-COMMERCE SITE WITH IMAGE RETRIEVAL

This application uses image processing to aid shopping. The developed system allows buyers to submit an image and will return similar images using the image retrieval techniques in the database. In general, our thesis includes two main sections: first, an e-commerce interface and second, the system database, which only the administrators of the system have the authority to access, manage, and modify by adding products or any other updates.

Through the first section, which is the UI, one can visit the webpage and see the entire product. The top part includes a header that displays the website phone number, languages, and currency. On the left side, many pages are listed (i.e., ladies, kids, sports, and technology) and can be opened with only a click.

The middle part contains a slideshow that presents new products. The most important part is the bottom, which has a product search engine where users can search for any product by uploading an image.

The second section holds the database of our thesis where the administrators can update the information and add products. The information in this section includes product details, such as name, size, etc. The bottom part is used for displaying the latest added products, which administrators have the authority to access and modify.

5.1. Main Page

On this page, buyers see the Holly website that contains in four types of products (ladies, kids, sports, and technology); each section has many products, as shown in Figure 5.1.

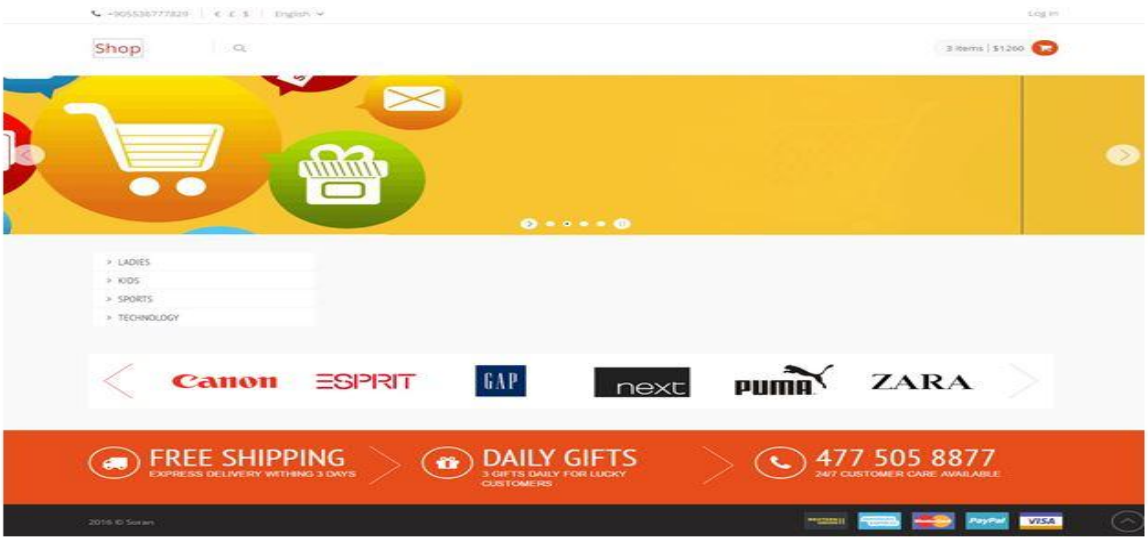

Figure 5.1. The Main Page of the System

Figure 5.2 shows an example of image retrieval for online shopping when the buyer wants to find an item; we submitted an image in the database and then searched in the website directly to get the result. Each product has an image, category, color, quantity, size, currency, and price of the product, which is included with each item.

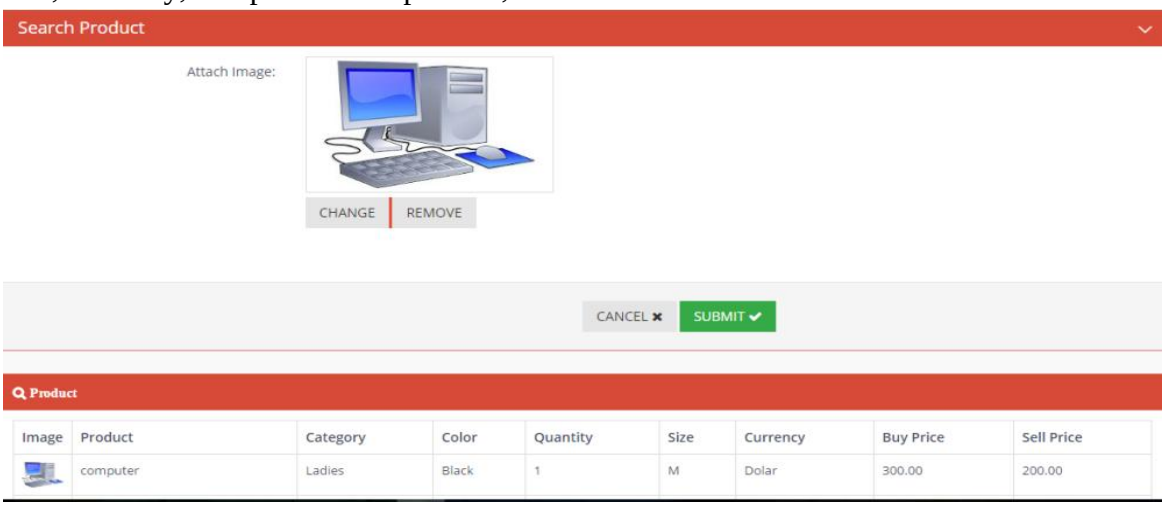

Figure 5.2. Product Retrieval by Image Search

The retrieval results provide the most similar images in the database directly to the website and show the results at the same time. On the other hand, this work has made the system convenient to the user, and it does not waste too much time searching for products. It is only necessary to upload any image of an item, and then the results will quickly display similar images, as the system has a fast connection to the database.

Figure 5.3 as displayed the result is more than one image; if we have multiple images in the database, they are shown in the result. This is better for the buyer to see many same or similar items. The main results in this thesis are represented by the retrieved image display accuracy. The retrieval results are not only one image but could be a group of images displayed by their similarity to the image that exists in the database.

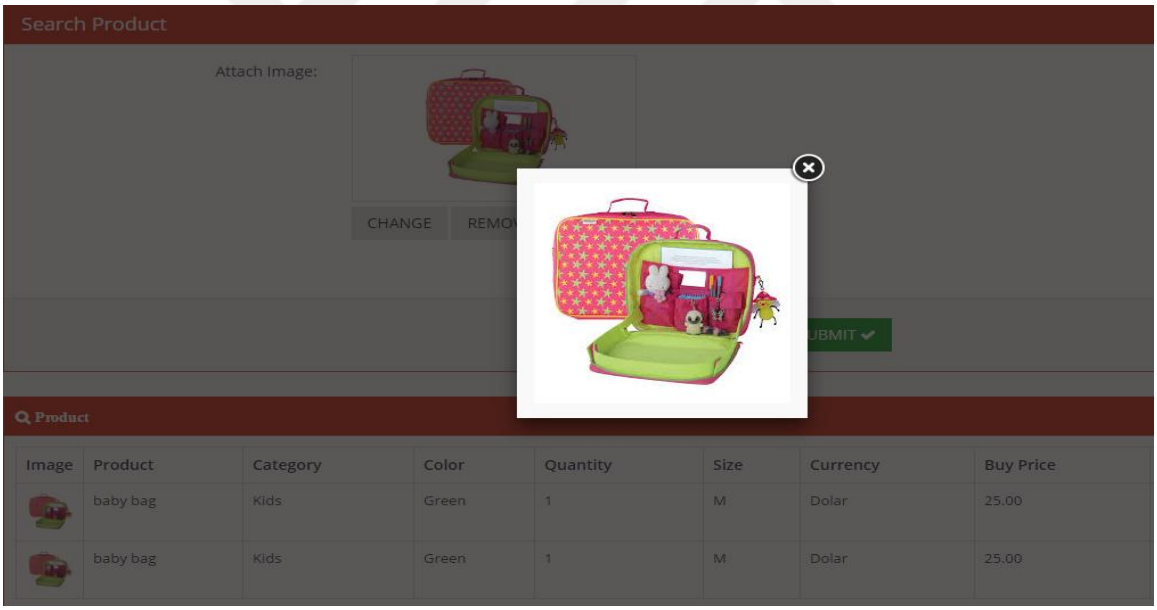

Figure 5.3. Image Search Multi-Result

Codes:

```
<i class="fa">Search Product</i>
   </div>
     <div class="tools"><a href="javascript:;" class="collapse" data-
original-title="" title=""> </a></div>
  </div>
    <div class="portlet-body form" id="blockui_sample_4_portlet_body">
 <!-- BEGIN FORM-->
    @using (Html.BeginForm("SearchByImage", "Shop", FormMethod.Post, new
{@id = "submit_form", enctype = "multipart/form-data", @class = "form-
horizontal"}))
{
<div class="form-body">  <div class="form-group ">
    label class="control-label col-md-3">Attach Image:</label>
<div class="col-md-6">
    <div class="fileinput fileinput-new" data-provides="fileinput">
    <div class="fileinput-preview thumbnail" data-trigger="fileinput"
style="width: 250px; height: 150px;"> </div>
 <div>
<span class="btn default btn-file">
<span class="fileinput-new"> Select file</span>
<span class="fileinput-exists green"> Change</span>
<input type="file" name="Image" id="Image" class="form-control" />
</span>
 <a href="#" class="btn default fileinput-exists" data-dismiss="fileinput">
Remove
</a><span for="picture" class="help-block
valid"></span></div></div></div></div>
</div>
<div class="form-actions fluid ">
<div class="row">
<div class="col-md-offset-6">
<a type="button" href="@Request.Url.AbsoluteUri" id="resetBtn" class="btn
default">Cancel <i class="fa fa-times"></i></a>
```

```
<button type="submit" class="btn green">Submit <i class="fa fa-
check"></i></button>
</div></div></div>}
<div class="portlet-body" id="blockui_sample_4_portlet_body">
<table class="table table-bordered" id="table1">
  <thead>
  <tr>
      <th>Image</th>
      <th>Product</th>
      <th>Category </th>
      <th>Color </th>
      <th>Quantity </th>
      <th>Size </th>
      <th>Currency </th>
      <th>Buy Price </th>
      <th>Sell Price </th>
    </tr>
  </thead>
```

The major function of the above-mentioned code is searching for a product by uploading an image by clicking on the select image button then submit it by clicking on the submit button and later we can see the result which includes the price, color, price …. Etc. and before submitting desired product for searching there are two other buttons, one for change the product and the other one is for canceling the whole uploading and searching process.

Figure 5.4 shows source image RGB components and destination image RGB components. The first one (a) is the source and the second one (b) is the destination or the results. It means the process for search image in our thesis is done in the destination.

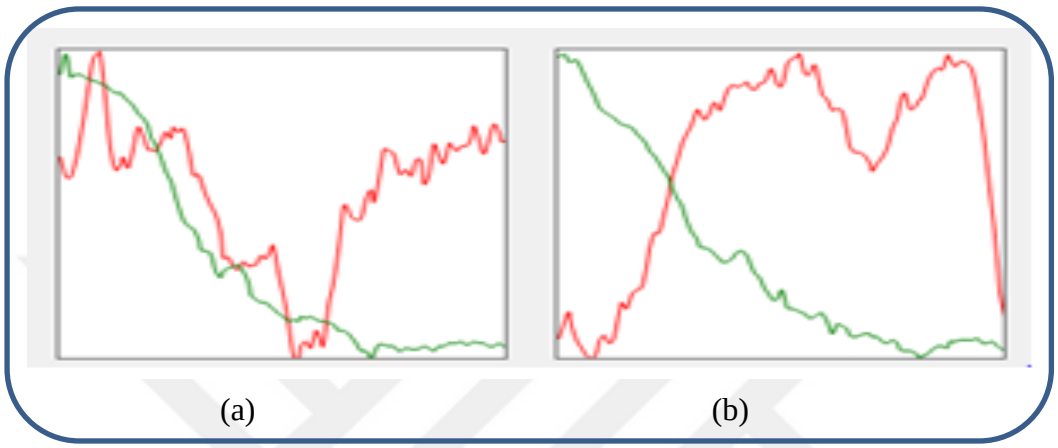

Figure 5.4. The Result of Image Search from Source to Destination

If our algorithm does not find a suitable result, it suggests the closest product according to color information in Figure 5.5 because this thesis, searches for images using RGB color to find the products. First, the user delivers the sample image by query interface; then, the system computes similar measurements according to the image features. Lastly, the system returns the query results according to the rank of similarity.

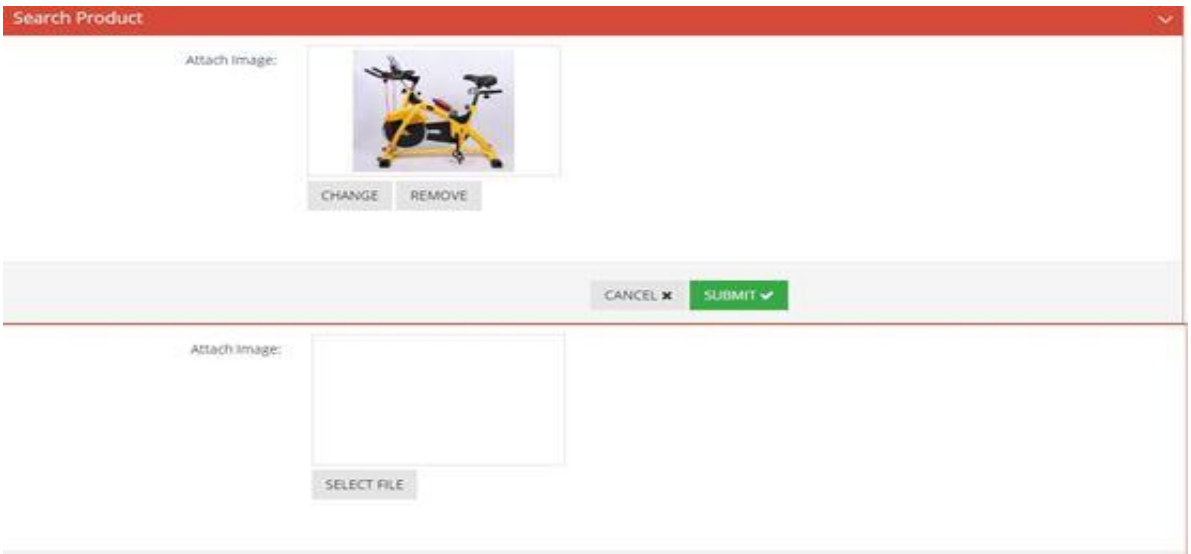

Figure 5.5. Suggest Retrieval Product

5.2. Add Items

Product information is added for new products. We use the add item buttons to upload the product, attach the image, and fill in the blanks. The blanks contain everything that is necessary for the customer. Then, we click on the submit button and the item will be added to the product list, as shown in Figure 5.6.

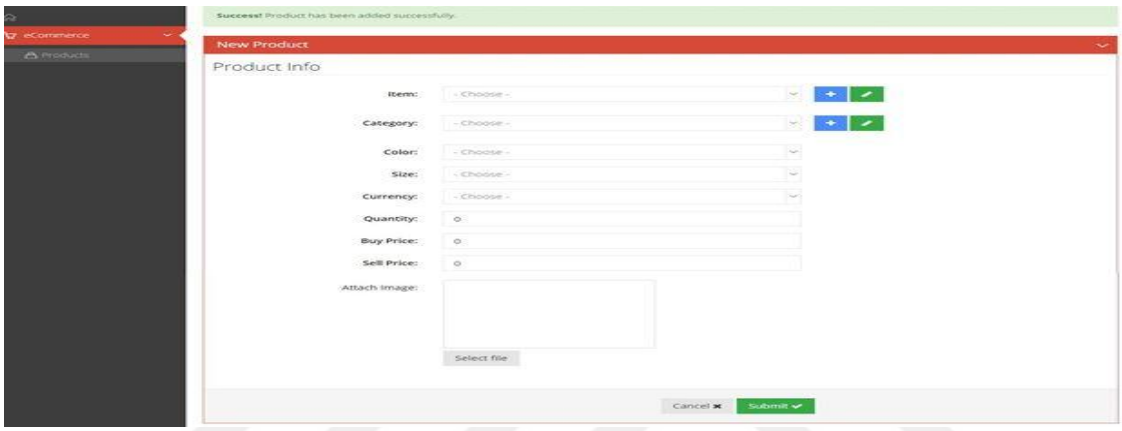

Figure 5.6. Add Products with Image in the Database

Figure 5.7 includes all the products that we will put in the thesis system. Only the authorized person can add products, monitoring the system, and update. Afterwards, the users or buyers can see the items on the website; each product has some specifications, such as category, color, quantity, size, currency, buy, and sell, and these will help customers find the exact item they desire. In this figure, products can be seen in series, meaning that we can choose to see five products at a time or more from the entry list.

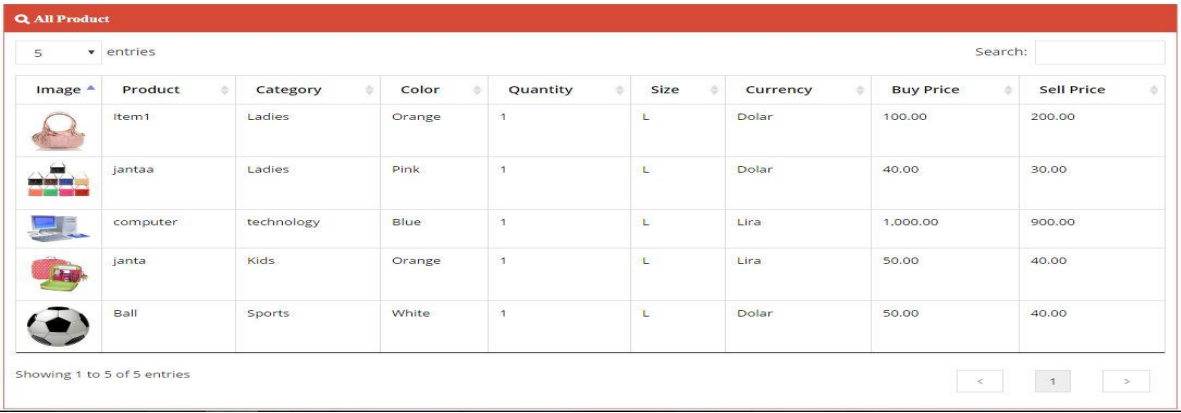

Figure 5.7. All products

In this thesis, the user can find products using the RGB method; the RGB connection index of the query image and the retrieved image can effectively retrieve images. The RGB connection index method is used to make a convenient retrieval decision. In addition, this is a new image retrieval search engine model, which has the best precision rate for great detection. However, we have many methods to find products using image retrieval, such as shape, texture, etc.

Experiments

The retrieved image was arranged by *Rindex* values in order *I*, and the first order is the maximum *Rindex* value and image similarity for the query image. The system performance measurement is based on the precision and recall. The recall rate of an image indicates the percentage of the relevant image retrieved. This parameter is defined as follows:

$$Recall = \frac{\text{Total number of Retrieval Relevant image}}{\text{Total number of relevant image}} \quad (5.1)$$

The precision rate of an image presents the average rank of the retrieved relevant image. This parameter is defined as follows:

$$Precision = \frac{\text{Total number of Retrieval Relevant image}}{\text{Total number of Retrieval image}} \quad (5.2)$$

Where *recall* is the ranking of the relevant image in the retrieval results; the *precision* 1 indicates the best retrieval and 0 indicates the worst retrieval [96].

The performance precision and recall of the relation method were calculated and used the same image database as Lin et al [97], Jhanwar et al. [98] and Huang Dai [99] and M.E. ElAlami [100]. Images for research purposes are employed, and the proposed results are compared with the results of [97-100]. The database consists of 100 images. These imaged are grouped into four types of product clusters, as shown in Table 5.1. This experiment used each image in each class as a query image. The precision P of each query image and finally obtain the average precision p/100 (100 images of a class. The experiment results from RGB relation index and the other four methods are shown in Table 5.2. The average

precision was presented in Table 5.2. It is obvious that the RGB relation method has achieved a batter average precision of various images than the other four methods. The precision performance of RGB relation method is 100 %.All corresponding images were correct to all query images.

Table 5.1. Sample of the Images in Four Clusters

Clusters	**Product Name**	**Sample of Images in the Each Cluster**
1	Ladies	
2	Kids	
3	Sports	
4	Technology	

Table 5.2. The Average Precision of these Methods

Product Name our Method	**Semantic Name other Methods**	**Present Method**	**Chuenet al.[97]**	**Jhanwar et al. [98]**	**Hung,Dai's [99]**	**M.E.ElAami [100]**
Ladies	Flowers	1.00	0.8910	0.8515	0.8975	0.914
Kids	Food	1.00	0.7325	0.3695	0.4265	0.741
Sports	Busses	1.00	0.8880	0.7410	0.8515	0.876
Technology	Horses	1.00	0.8025	0.5680	0.5890	0.834
Average		1.00	0.8285	0.6325	0.6911	0.841

The system has been developed based on the experience gained through our investigation of other systems with similar methods to this system. It is important to mention the following drawbacks of another thesis that differs from this thesis: slow image search in e-commerce, lack of retrieving identical or desired results in many cases,

retrieving single examples when a product is searched for, and inaccuracy. The researchers of this study have endeavored to overcome these problems or at least mitigate them [101-103].

In this thesis, we have attempted to solve all these problems with improved speed. As a result, user satisfaction can be achieved when users find the products easily. This approach leads in showing identical items in the database and suggesting them to the users. Searching for product information by keywords has some issues, including more search area, inaccuracy in search results, and indirect search [104]. In this thesis, to solve these problems, similar image retrieval is a new vision in image searching technology. This technology searches images according to their content and context environment. First, some objective image features are extracted automatically by the computer, such as color. At that point, the image is recorded in the database in view of these components. According to the sample image, the users or buyers recover the image in the database using some similarity estimations, and the framework returns comparative outcomes.

This technique provides another search mode – searching by image. Searching using an image will help clients locate the same or comparable images from many web images. It relieves the frustration of buyers when they search for an image that is hard to portray by keywords, and they do not need to search by keywords. The visualization of searching by image fulfills the clients' need for higher searching prerequisites; helps clients gain item information, and elevates web-based shopping knowledge. This system connects the information platform with the purchase platform. The customers can automatically search for similar images according to their favorite images from the information platform using this system. Eventually, it will increase the purchase ratio of customers. The system can also automatically collect the product information from the other online shopping platforms, provide direct similar image retrieval for customers, and display the search results with a customized way. The vision image searching outcomes without other input can help clients purchase the item.

The Most e-commerce websites [105] use text to give information about products. There is another point that separates this research from others, which is that searching for products can be done primarily through text, and this procedure has a disadvantage which is that the text on the website will be a problem for some other nations or uneducated persons without the ability of writing.

Although this represents a very useful technique, it must be described by humans. Manual labeling for products is very costly, time-consuming, and infeasible for many applications. It requires a large amount of human labor to manually annotate large-scale image databases.

We can see in the [106], the mentioned website was created using PHP, but in this thesis, we used C#, as we know C# is one of the best programming languages that can be used in image processing.

In this thesis, similar image retrieval in e-commerce for online shopping is based on color and edge directivity, and the specific important points are listed as follows:

1. Search speed and efficiency: This thesis presents an active retrieval scheme. Since the image database has many images, simultaneous image retrieval is needed.
2. Search accuracy: In this thesis, the process of image retrieval in the database is produced and shown exactly.
3. Data security and accuracy are improved.
4. In this study, the user can use the image to help in shopping.
5. This method permits the users to present an image and will return a list of comparable items utilizing image retrieval strategies.
6. This helps users investigate their alternatives and choose precisely what they need.

6. CONCLUSIONS

We live in a world full of technological applications that are developed every day. In the modern hi-tech world, there is a progressive elimination of the differences between psyche and the environment, between waking and dreaming, and between the conscious and the subconscious. In recent years, an image search for finding products appeared as an important research field. This research combines both online shopping and image retrieval to help users choose the best product. This work has improved the implementation of image search engines by designing a method that re-arranges the original results of any query. The purpose of re-arranging is to show the most similar images. For this study, the users are provided with an online shopping website that can be used to buy goods. To implement this application, ASP.NET was used as the technology. In addition, have used MYSQL as the back-end database, as it is one of the most common open-source databases, providing secure, quick data archival, and simplicity.

This work also takes advantages of similar image retrieval for online shopping, which is an image getting in a computer system for searching and retrieving image from a huge database that computerizes image and it helps the user to choose the best type of products to submit image search engine on the website. We think that the suggested technique may help to decrease the semantic spread between high-level idea and low-level features. The results of this work enhance the understanding of different techniques for feature extraction and similarity measurement, which assist in image recovery in e-commerce. In the thesis, has mentioned that the image search engine is better than a textual search because we know the users can easily find products while online shopping to get what they want.

In the future, we aim to improve the current system, fix all bugs, and test more datasets, such as oracle, to compare the results with the previous database. Due to the prompt response of the system, more users can access the system at the same time. In addition, more search buttons will be added so that more search options will be available. In addition, we plan to develop a technique for suitable product retrieval for the shopper or user and to solicit feedback from the users can help us to determine the best user requirements.

REFERENCES

[1] **Ling, K. C., Chai, L. T., & Piew, T. H**., (2010), "The effects of shopping orientations, online trust and prior online purchase experience toward customers' online purchase intention", *International Business Research*, *3*(3), 63.

[2] **BULUT, Z. A**., (2015), "Determinants of Repurchase Intention in Online Shopping: A Turkish Consumer's Perspective", *International Journal of Business and Social Science*, *6*(10), 55-63.

[3] **Hsu, E., Paz, C., & Shen, S**., (2011), "Clothing image retrieval for smarter shopping", *California*.

[4] **Kim, S.Kyun**, (2016), "Query based image retrieval and object extraction applied in e-commerce service," Computer engineering department, Myongijy University, Korea.

[5] **Iakovidou, C., Anagnostopoulos, N., Kapoutsis, A. C., Boutalis, Y., & Chatzichristofis, S. A**., (2014, June), "Searching images with MPEG-7 (& mpeg-7-like) powered localized descriptors: the SIMPLE answer to effective content-based image retrieval", In *Content-Based Multimedia Indexing (CBMI), 2014 12th International Workshop on* (pp. 1-6). IEEE.

[6] **Chen, M., Chen, A. N., & Shao, B. B**., (2003), "The implications and impacts of web services to electronic commerce research and practices", *J. Electron. Commerce Res.*, *4*(4), 128-139.

[7] **Nah, F. F. H., & Davis, S**., (2002), "HCI research issues in e-commerce", *Journal of Electronic Commerce Research*, *3*(3), 98-113.

[8] **Zwass, V**., (1999), "Structure and macro-level impacts of electronic commerce", *Emerging Information Technologies: Improving Decisions, Cooperation, and Infrastructure, Sage, Beverly Hills, CA*, 289-315.

[9] **Spink, A., & Jansen, B. J**., (2008), "Trends in searching for commerce related information on web search engines", *Journal of Electronic Commerce research*, *9*(2), 154.

[10] **Goel, N., & Sehgal, P**., (2014), "Image Retrieval Using Fuzzy Color Histogram and Fuzzy String Matching: A Correlation-Based Scheme to Reduce the Semantic Gap", In *Intelligent Computing, Networking, and Informatics* (pp. 327-341). Springer India.

[11] **Sherman, C., & Price, G**., (2003), "The invisible web: uncovering sources search engines can't see", *Library Trends*, *52*(2), 282-298.

[12] **Zahavy, T., Magnani, A., Krishnan, A., & Mannor, S**., (2016), "Is a picture worth a thousand words? A Deep Multi-Modal Fusion Architecture for Product Classification in e-commerce", *arXiv preprint arXiv:1611.09534*.

[13] **Datta, R., Joshi, D., Li, J., & Wang, J. Z**., (2008), "Image retrieval: Ideas, influences, and trends of the new age", *ACM Computing Surveys (Csur)*, *40*(2), 5.

[14] **Deselaers, T., Keysers, D., & Ney, H**., (2008), "Features for image retrieval: an experime".

[15] **Penatti, O. A. B., & da Silva Torres, R**., (2008, October), "Color descriptors for web image retrieval: a comparative study", In *Computer Graphics and Image Processing, 2008. SIBGRAPI'08. XXI Brazilian Symposium on* (pp. 163-170). IEEE.

[16] **Annesley, J., Orwell, J., & Renno, J. P**., (2005, October), "Evaluation of MPEG7 color descriptors for visual surveillance retrieval", In *Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005. 2nd Joint IEEE International Workshop on* (pp. 105-112). IEEE.

[17] **Van De Sande, K., Gevers, T., & Snoek, C**., (2010), "Evaluating color descriptors for object and scene recognition", *IEEE transactions on pattern analysis and machine intelligence*, *32*(9), 1582-1596.

[18] **Loncaric, S**., (1998), "A survey of shape analysis techniques", *Pattern Recognition*, *31*(8), 983-1001.

[19] **Zhang, D., & Lu, G**., (2004), "Review of shape representation and description techniques", *Pattern Recognition*, *37*(1), 1-19.

[20] **Safar, M., Shahabi, C., & Sun, X**., (2000), "Image retrieval by shape: a comparative study", In *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on* (Vol. 1, pp. 141-144). IEEE.

[21] **Xu, F., & Zhang, Y. J**., (2006), "Evaluation and comparison of texture descriptors proposed in MPEG-7", *Journal of Visual Communication and Image Representation*, *17*(4), 701-716.

[22] **Howarth, P., & Rüger, S.,** (2004, July), "Evaluation of texture features for content-based image retrieval" ,*In International conference on image and video retrieval* (pp.

326-334). Springer Berlin Heidelberg.

[23] **Spyrou, E., Le Borgne, H., Mailis, T., Cooke, E., Avrithis, Y., & O'Connor, N.,** (2005), "Fusing MPEG-7 visual descriptors for image classification", Artificial Neural Networks: Formal Models and Their Applications–*ICANN* 2005, 747-747.

[24] **Bama, B. S., Valli, S. M., Raju, S., & Kumar, V. A.,** (2011), "Content-based leaf image retrieval (CBLIR) using shape, color and texture features", *Indian Journal of Computer Science and Engineering*, 2(2), 202-211..

[25] **Vijayasarathy, L. R., & Jones, J. M.,** (2000), "Print and Internet catalog shopping: assessing attitudes and intentions", *Internet Research*, 10(3), 191-202.

[26] **Phau, I., & Meng Poon, S.,** (2000), "Factors influencing the types of products and services purchased over the Internet", *Internet Research*, 10(2), 102-113.

[27] **Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., ... & Steele, D.,** (1995), "Query by image and video content: The QBIC system", *Computer*, 28(9), 23-32.

[28] **Wang, J. Z., Li, J., & Wiederhold, G.,** (2001), "SIMPLIcity: Semantics-sensitive integrated matching for picture libraries", *IEEE Transactions on pattern analysis and machine intelligence*, 23(9), 947-963.

[29] **Wong, K. M., Cheung, K. W., & Po, L. M.,** (2005, May), "MIRROR: an interactive content based image retrieval system",*In Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium* on (pp. 1541-1544). IEEE.

[30] **Martinez, J. M., & Le Toan, T.** (2007). Mapping of flood dynamics and spatial distribution of vegetation in the Amazon floodplain using multitemporal SAR data. Remote Sensing of Environment, 108(3), 209-223.

[31] **Manjunath, B. S., Ohm, J. R., Vasudevan, V. V., & Yamada, A.,** (2001), "Color and texture descriptors", *IEEE Transactions on circuits and systems for video technology*, 11(6), 703-715.

[32] **Mussarat, Y., Muhammad, S., Sajjad, M., & Isma, I.,** (2013), "Content-based image retrieval using combined features of shape, color and relevance feedback", *KSII Transactions on Internet and information systems*, 7(12), 3149-3165.

[33] **Rosen, A.,** (2002), "The e-commerce question and answer book: a survival guide for

business managers", *AMACOM Div American Mgmt Assn.*

[34] **Demaray, M. K., & Malecki, C. K.,** (2002), " The relationship between perceived social support and maladjustment for students at risk", *Psychology in the Schools*, 39(3), 305-316.

[35] **Mesenbourg, T.L.**,(2000),"Measuring Electronic Business: Definitions", Underlying Concepts, and Measurement Plans.

[36] **Lallana, E., Quimbo, R., & Andam, Z. R.,** (2000), "An Introduction to eCommerce: definition adapted and expanded",Philippines: *DAI-AGILE*, 17.

[37] **Varian, H. R.,** (1999), "Market structure in the network age", *Understanding the Digital Economy, MIT Press*, Cambridge, MA, 137-150.

[38] **Fang, Y., Qureshi, I., Sun, H., McCole, P., Ramsey, E., & Lim, K. H.,** (2014), "Trust, Satisfaction, and Online Repurchase Intention: The Moderating Role of Perceived Effectiveness of E-Commerce Institutional Mechanisms", Mis Quarterly, 38(2), 407-427.

[39] **Rayport, J. F., & Jaworski, B. J.**, (2002), "Introduction to e-commerce", McGraw-Hill/Irwin marketspaceU.

[40] **S.Kadam, P.Gupta, C.Veel, A.Loke**, (2016), "Information technology", Vidyalankar Institute of Technology, Mumbai, India, *International Journal of Advanced Research in Computer and Communicate Engineering*, Vol.5, Issue 33.

[41] **Goswami, A., Chittar, N., & Sung, C. H.,** (2011, March), "A study on the impact of product images on user clicks for online shopping", *In Proceedings of the 20th international conference companion on World wide web* (pp. 45-46). ACM.

[42] **Kasemsap, K.,** (2015), "The role of electronic commerce in the global business environments", Handbook of research on interactive information quality in expanding social network communications, 304-324.

[43] **Kalakota, R., & Robinson, M.,** (2003), "Services Blueprint: Roadmap for Execution", Addison-Wesley Professional.

[44] **Mesenbourg, Thomas L.,** (2001), "Measuring electronic business", *US Census Bureau*, http://www. census. gov/stats .

[45] **Sachs, G.,** (1999), "E-Commerce/Internet", *Goldmann Sachs Investment Research*

[46] **Andam, Z. R.**, (2003), "E-commerce and e-business, e-primer for the information economy", *Society and Polity Series, UNDP-APDIP and e-ASEAN Task Force*.

[47] **Kalakota, R., & Whinston, A. B.,** (1997), "Electronic commerce: a manager's guide", *Addison-Wesley Professional.*

[48] **Lallana, E. C., Pascual, P., & Andam, Z. R.**, (2002), "SMEs and e-COMMERCE In Three Philippine Cities", *The Asia Foundation. Digital Philippines*.

[49] **Mokhtarian, P. L.,** (2004), "A conceptual analysis of the transportation impacts of B2C e-commerce", *Transportation*, 31(3), 257-284.

[50] **Ho, C.Fu, & W.Wu,** (1999), "Antecedents of customer satisfaction on the Internet: an empirical study of online shopping", Systems Sciences HICSS-32. *Proceedings of the 32nd Annual Hawaii International Conference on. IEEE.*

[51] **Wood & M. Charles,** (2004), "Marketing and e-commerce as tools of development in the Asia-Pacific region: a dual path", *International Marketing Review* 21.3: 301-320.

[52] **Sharma and Dhiraj,** (2002), "Role of Electronic Commence in Business Environment* Tigist Wolde Gebre".

[53] **Levy, Phil.,** (2012), "Simple Strategies for Enhancing eCommerce Profitability", *A First Data White Paper.*

[54] **Siwicki, B.,** (2011), "20% of e-mail sent by retailers is opened on a mobile device",*Knotice study reports*. Retrieved June, 10, 2011.

[55] **Jones, Matthew T.,** (2012), "The PayPal Official Insider Guide to Mobile Profits: Make Money Anytime, Anywhere", PayPal Press.

[56] **Lin, X., Gokturk, B., Sumengen, B., & Vu, D.,** (2008, January), "Visual search engine for product images", *In Electronic Imaging 2008 (pp. 68200M-68200M). International Society for Optics and Photonics.*

[57] **Niranjanamurthy, M., Kavyashree, N., Jagannath, S., & Chahar, D.,** (2013), "Analysis of e-commerce and m-commerce: advantages, limitations and security issues". *International Journal of Advanced Research in Computer and Communication Engineering*, 2(6).

[58] **Datta, R., Joshi, D., Li, J., & Wang, J. Z.,** (2008), "Image retrieval: Ideas, influences, and trends of the new age", *ACM Computing Surveys (Csur),* 40(2), 5.

[59] **Riad, A. M., Atwan, A., El-Bakry, H. M., EL-Ghany, S. A., & Mastorakis, N.**, (2009), "An Intelligent Distributed Algorithm for Efficient Web Image Retrieval", *International Journal of Communication*, 63-76.

[60] **Zhang, J., Liu, X., Luo, J., & Lang, B.,** (2010, December), "Dirs: Distributed image retrieval system based on mapreduce", *In Pervasive Computing and Applications (ICPCA), 2010 5th International Conference on (pp. 93-98). IEEE.*

[61] **Zhou, D., Hu, B., Wang, Q., Hu, B., Jia, L., Wu, Y., & Xie, L.,** (2013, April), "Design of Shopping Guide System with Image Retrieval Based on Mobile Platform", *In 2nd International Symposium on Computer, Communication, Control and Automation.* Atlantis Press.

[62] **Kruliš, M., Skopal, T., Lokoč, J., & Beecks, C.,** (2012),"Combining CPU and GPU architectures for fast similarity search", Distributed and Parallel Databases, 30(3-4), 179-207.

[63] **Pavlidis, T.,** (2008, December), "Limitations of content-based image retrieval", In Invited Plenary Talk at the *19th Internat. Conf. on Pattern Recognition, Tampa, Florida,* December (pp. 8-11).

[64] **Srinivasan, S. S., Anderson, R., & Ponnavolu, K**., (2002), "Customer loyalty in e-commerce: an exploration of its antecedents and consequences", *Journal of Retailing*, *78*(1), 41-50.

[65] **Fingar, P., Kumar, H., & Sharma, T.,** (1999), "Enterprise e-commerce," *Meghan-Kiffer Press.*

[66] **Aldrich, D. F.,** (1999), "Mastering the digital marketplace: practical strategies for competitiveness in the new economy", John Wiley & Sons.

[67] **Bakos, J. Y.,** (1997), "Reducing buyer search costs: Implications for electronic marketplaces", *Management Science*, 43(12), 1676-1692.

[68] **Lynch Jr, J. G., & Ariely, D.,** (2000), "Wine online: Search costs affect competition on price, quality, and distribution", *Marketing Science*, 19(1), 83-103.

[69] **Windham, L., & Orton, K.**, (2000), "The soul of the new consumer: The attitudes, behavior, and preferences of e-customers", Allworth Press.

[70] **Seybold, P. B., & Marshak, R. T.,** (1999), "Customers. com: how to create a profitable business strategy for the Internet and beyond", Random House Audio Assets.

[71] **Bakos, Y.**, (1998), "The emerging role of electronic marketplaces on the Internet", *Communications of the ACM*, 41(8), 35-42.

[72] **Kherfi, M. L., Ziou, D., & Bernardi, A.,** (2004), "Image retrieval from the world wide web: Issues, techniques, and systems", *ACM Computing Surveys (CSUR)*, 36(1), 35-67.

[73] **El Gawady, Z. M.**, (2005, November), "The impact of e-commerce on developed and developing countries case study: Egypt and united states", *In International Conference of Globalization. Technology and Sustainable Development.*

[74] **Barndorff-Nielsen, O.,** (2014), "Information and exponential families in statistical theory", John Wiley & Sons.

[75] **Aydın, E., & Savrul, B. K.,** (2014), "The Relationship between Globalization and E-Commerce: Turkish Case", *Procedia-Social and Behavioral Sciences*, 150, 1267-1276.

[76] **Long, B., Bian, J., Dong, A., & Chang, Y.,** (2012, October), "Enhancing product search by best-selling prediction in e-commerce", *In Proceedings of the 21st ACM international conference on Information and knowledge management (pp. 2479-2482). ACM.*

[77] **Beall, J.,** (2008), "The weaknesses of full-text searching", *The Journal of Academic Librarianship*, 34(5), 438-444.

[78] **Gevers, T., & Smeulders, A. W. M.,** (2004), "Image search engines: An overview", *Emerging Topics in Computer Vision*, 1-54.

[79] **Wang, J. Z., Wiederhold, G., Firschein, O., & Sha, X. W.**, (1998), "Content-Based Image Indexing and Searching Using Daubechies' Wavelets", *In Digital Libraries* (Vol. 1, No. 4, pp. 311-328).

[80] **Keijzers, S., Maandag, P., Marchiori, E., & Sprinkhuizen-Kuyper, I. G.,** (2013), "Image Similarity Search using a Negative Selection Algorithm", *In ECAL* (pp. 838-845).

[81] **Alzoubi, A. H.,** (2015), "Comparative analysis of image search algorithm using average RGB, local color histogram, global color histogram and color moment HSV", (Doctoral dissertation, *Universiti Tun Hussein Onn Malaysia*).

[82] **J. Parag, S.S.Prakash, P.Mihir, an& d V. Omkar**, (2015)," Image Search Engine Using SIFT Algorithm", *Journal of The International Association of Advanced Technology and Science*.

[83] **Mathew, S. P., Balas, V. E., Zachariah, K. P., & Samuel, P.,** (2014), "A content-based image retrieval system based on polar raster edge sampling signature", *Acta Polytechnica Hungarica*, 11(3), 25-36.

[84] **Zhang, J., & Zou, W.,** (2010, March), "Content-based image retrieval using color and edge direction features", *In Advanced Computer Control (ICACC), 2010 2nd International Conference on (Vol. 5, pp. 459-462). IEEE.*

[85] **Solli, M., & Lenz, R.,** (2008, January), "Color emotions for image classification and retrieval", *In Conference on Colour in Graphics, Imaging, and Vision (Vol. 2008, No. 1, pp. 367-371). Society for Imaging Science and Technology.*

[86] **Hu, Y., Yin, H., Han, D., & Yu, F.,** (2014), "The application of similar image retrieval in electronic commerce", *The Scientific World Journal*, 2014.

[87] **Singh, S. M., & Hemachandran, K.**, (2012), "Content-based image retrieval based on the integration of color histogram, color moment and Gabor texture", *International Journal of Computer Applications*, 59(17).

[88] **Kinnaree, P., Pattanasethanon, S., Thanaputtiwirot, S., & Boontho, S.,** (2011), "RGB color correlation index for image retrieval", *Procedia Engineering*, 8, 36-41.

[89] **Chaudhari, Reshma, & Patil, A. M.**, (2012),"Content-based image retrieval using color and shape features", *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* 1.5: 67-72.

[90] **Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P.**, (2004), "Image quality assessment: from error visibility to structural similarity", *IEEE Transactions on image processing*, *13*(4), 600-612.

[91] **Lucassen, John M., and Stephane H. Maes.,** (Feb. 2006), "MVC (model-view-controller) based multi-modal authoring tool and development environment", *U.S. Patent* No. 6,996,800. 7.

[92] **Shamoi, P., Inoue, A., & Kawanaka, H.**, (2015, July), "Deep color semantics for E-commerce content-based image retrieval", *In Proceedings of the 2015 International Conference on Fuzzy Logic in Artificial Intelligence-Volume* 1424 (pp. 14-20). *CEUR-WS. org*

[93] **Halvorson, M.,** (2013), "Microsoft visual basic 2013 step by step. Pearson Education".

[94] **Sempf, B., Sphar, C., & Davis, S. R.**, (2013), "*C# 5.0 All-in-one for Dummies*", John Wiley & Son.

[95] **Mistry, R., & Misner, S.**, (2014), "Introducing Microsoft SQL Server 2014", *Microsoft Press.*

[96] **K.Zeng, N.Wu, and K.K. Yen,** (2014), "A color boosted local feature extraction method for mobile product search", *International Journal on Recent Trends in Engineering & Technology* 10.2: 78.

[97] **Lin, C. H., Chen, R. T., & Chan, Y. K.,** (2009), "A smart content-based image retrieval system based on color and texture feature", *Image and Vision Computing*, 27(6), 658-665.

[98] **Jhanwar, N., Chaudhuri, S., Seetharaman, G., & Zavidovique, B.**, (2004), "Content-based image retrieval using motif cooccurrence matrix", *Image and Vision Computing*, 22(14), 1211-1220.

[99] **Huang, P. W., & Dai, S. K.,** (2003), "Image retrieval by texture similarity", *Pattern Recognition*, *36*(3), 665-679.

[100] **ElAlami, M. E.,** (2011), "A novel image retrieval model based on the most relevant features", *Knowledge-Based Systems*, 24(1), 23-32.

[101] **Zuva T. & Ngwira**, **S. M.,** (July 2012), "IMAGE content in location-based shopping recommender systems for mobile users ", (Department of computer engineering ,Tshwane University of Technology ,Pretoria ,South Africa) , Oludayo O. Olugbara(Department of information technology ,Durban University of technology , Durban,South Africa) and Sunday O. Ojo (Faculty of information and communication technology ,Tshwane University of Technology ,Pretoria ,South Africa).. Advanced Computing: *An International Journal (ACIJ)*, vol.3.No.4).

[102] **Lin, K., Yang, H. F., Liu, K. H., Hsiao, J. H., & Chen, C. S.,** (2015, June), "Rapid clothing retrieval via deep learning of binary codes and hierarchical search",

In Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (pp. 499-502). ACM.

[103] **DOS SANTOS, C. P., FERNANDES, D., & ANPAD.**, (2004), "Failure and recovery in internet purchasing", *Anais do XXXI ENANPAD*, 22-26.

[104] **Hu, Y., Yin, H., Han, D., & Yu, F.**, (2014), "The application of similar image retrieval in electronic commerce",*The Scientific World Journal*, *2014*.

[105] **Shamoi, P., Inoue, A., & Kawanaka, H.,** (2015, July), "Deep color semantics for E-commerce content-based image retrieval", *In Proceedings of the 2015 International Conference on Fuzzy Logic in Artificial Intelligence-Volume* 1424 (pp. 14-20). *CEUR-WS. org.*

[106] **Liu, Y.**, (2016), "Intermediary System Using Image Classification for Online Shopping".

CURRICULUM VITAE (CV)

Personal Information

Full name: Soran Abdulkarim Pasha Al-Jaf

Place of birth: Kalar-Sulaymaneah-Iraq

Date of Birth: 10 / February / 1986

Marital status: Married

Contact Information

Phone: 00964 (0) 770 2297229

Address: Kalar-Sulaymaneah-Iraq

Email: Soranpasha134@gmail.com

Education

2007-2011 BSc. Computer Systems Engineering at Koya University, Iraq

2015-present MSc. student in Software Engineering at Firat University, Turkey