# Journal of Vibration Engineering

Registered

Our Website
www.jove.science

# A Review on Enhancing Retrieval Performance of Content Based Image Retrieval System Based on Early and Late Fusion Techniques

[1]PranavLod[2]Prof.V.S.Mahalle
[1]DepartmentofComputerScienceandEngineeringSSGMCE,Shegaon
[2]SantGadgeBabaAmravatiUniversity,Amravati,Maharashtra

## Abstract

The query image is used to get pictures from a database or image repository by Content-BasedImageRetrieval(CBIR)systems.CBIRiscrucialinanumberofindustries,includingarchitecture, web-based search, criminal prevention, and medical diagnostics.In CBIR, thereare primarily two stages: the first is feature extraction, and the second is similarity matching.The effectiveness and efficacy of CBIR may be increased in a number of ways, includingsegmentation, relevance feedback, expanding searches, and fusion-based techniques. Severaltechniquesformixingandfusingseveralpicturedescriptorshavebeenproposedintheliterature. Early and late fusion techniques make up the two main categories into which fusionstrategies are commonly separated. While late fusion refers to either the combining of outputsprovided by distinct retrieval methods or the combination of different rankings of similarity,early fusion is the integration of visual characteristics from several descriptors into a singlevectorpriorto thesimilarity computation.

**KeyWords:**Content-basedimageretrieval,Featureextraction,Fusionmethod.

## 1.     Introduction

Large image databases can be searched using one of two methods. The first method, text-basedimage retrieval (TBIR), uses manually annotated images that have appropriate names attachedto them. The end user then uses keywords to search for and find the required images. Themanual annotating process takes a large amount of time, which is the first obvious drawback ofthis strategy. The second drawback is the amount of onus placed on the end user to create theirown queries. Content-based image retrieval (CBIR), an alternative and efficient technique,offers several benefitsand addressesthe keyissues raised by the prior technique. Due tothefast growth of multimedia data through contemporary sources including the Internet, smartdevices, internet of things (IoT) devices and sources, social networks, and medical imagingsources during the past two decades, CBIR also became the primary approach and a focus ofextensive study. However, single feature spaces are not always enough to produce the greatestretrieval results. The majority of the older CBIR research were based on individual or singlefeature descriptors retrieved from colour, texture, or form contents/descriptors. As a result,

themajorityofcurrentresearchhasincorporatedmanydescriptorsfromtheaforementionedinstances. The results of an early survey of CBIR's vast and varied application areas werepresented.

Whilelatefusionreferstoeitherthecombiningofoutputsprovidedbydistinctretrievalmethods or the combination of different rankings of similarity, early fusion is the integration ofvisualcharacteristicsfromseveraldescriptorsintoasinglevectorpriortothesimilaritycomputation. In this study, it is suggested that a set of colour and texture attributes be employedforbothfusion procedures.

## 2.      Reviewofliterature

[1] Ali Ahmed et al. suggested a retrieval technique that makes use of colour and linguisticinformation, implements and contrasts the two popular forms of fusion, and then recommendsthe optimal strategy to recover related photos. The following are the primary contributions ofourwork:

• For extracting comparable pictures based on the fusion of visual and linguistic properties, acompletely automated CBIR technique is given.When high-quality representative extractedfractures of both colour and text-based pictures are concatenated into a single vector, thesuggestedtechniqueprovides satisfactoryretrievaloutcomes.

• The system's performance retrieval results improve as the accuracy of the characteristics ituses increases. The technique also supports the concept of combining several results fromrelated measurements. Any combination of similarity measurements that the user chooses mayprovide an excellent final fused precision value that is superior to the sum of the separatesimilarityvalues.

Various methods to merge and fuse different picture descriptors have been suggested in theliterature throughout the years. Early and late fusion are the two categories into which fusionmethods are often divided. Five categories of fusion strategies have been developed: traditionalearly fusion, features reweighting for early fusion, representation by multi-feature spaces

forlatefusion,relevancefeedbackapproaches,andmultimodalretrieval.Thedefinitionandseveralrecentresearchthatusedtheseconceptsareprovidedinthesectionsthatfollow.Multiple        picture representations (such as colour, texture, and form) are integrated into a singlefeature vector in normal or classic early fusion. This method is clear-cut, extremely easy tounderstand, and often assigns an equivalent weight value scheme to any of the feature spaces.Many research have used this approach. Earlyfusion with feature weighting is appropriatewhentheretrieval process is [1].

[1] Ahmed A et al. noted that feature selection is an important pre-processing operation in datamining and machine learning processes; it has a favourable effect in terms of reducing theinformation repetition and high dimensionality of data. The practise of eliminating elementsfrom the data collection that are irrelevant to the task at hand is referred to as feature selection.The choice of features is important for several reasons, including performance, computationaleffectiveness,featureinterpretability,andsimplicity.Byminimisingunimportantandpotentiallyredundantfeatures,increasingprecision,reducingunpredictabilityandtheassociated computationalcost, and increasing likelihood, feature selection is used to achievetheshared

objectiveand commongoal ofimprovingaccuracy oftheclassifier[2].

[3,4] Malebary S et al. discovered that there are now more diverse digital picture databasesavailable. In response to customer feedback that searching for and recovering necessary photosfrom big collections is a highly challenging undertaking, successful and efficientretrievaltechniqueshavebeencreated.Numerousvisualqualitieshavebeenregardedasindirecttechniquesofretrievingpicturesfromdatabasesinthemajorityofcontent-basedimageretrieval(CBIR)systems[5,6].Amethodologyforobtainingmedicalpicturesbasedonautomatically deduced visual attributes, such as colour and texture, is known as content-basedmedical image retrieval (CBMIR). Retrieval performance remains one of the most difficultissues in current CBMIR studies, despite the many methods and approaches that have beenproposed. This is because of the well-known "semantic gap" problem, which occurs betweenmachine-captured low-level image features and human-perceived high-level semantic concepts.There has been a lot of planned research to close this gap. The retrievalmodel's precision canbe increased and improved using two expansion strategies presented inthis paper, both ofwhich rely on the top-ranked pictures. Our first expansion technique involves reformulating thenew expand query picture based on the average values for the characteristics of the top-rankedphotos, whereas our second expansion method just considers the most crucial aspects. Twelvetexture characteristics and eighteen colour features were derived from two widely used medicalpicturedatasets,AtreyPK,HossainMA,andElSaddik[7,8],andusedintheexpansionprocedure.

## 3.    ResearchMethodology

### 3.1. Processflow

- The image is preprocessed using image resizing, and the features are retrieved using theimage'scolourand texture.
- Theuser canchooseanimagefromthedataset.
- Thecomparablephotoswillthenbeobtainedoncethesimilarityisassessedtocategorisetheimage.
- Performance metrics like recall and accuracy will be estimated for the categorizationperformance.
- Early fusion is predicated on retrieving related photos as soon as the item is discovered.After locating the item, late fusion looks through the most pertinent photos and retrievesthem oneat atime.

## 4.    EarlyFusionTechniques

Unimodal characteristics are initially extracted using indexing strategies that rely onearlyfusion.Followingexaminationoftheseveralunimodalstreams,theretrievedcharacteristicsareintegratedintoasinglerepresentation.Earlyfusionapproachesusesupervised learning to categorise semantic ideas after combining unimodal information in amultimodal representation. Due to the fact that all of the characteristics are combined from thebeginning, early fusion can provide a genuinely multimodal feature representation. Early fusionalso offers the benefit of just requiring one learning step. The difficulty of fusing all the traitsintoasinglerepresentation

isthismethod's drawback.

## 5.        LateFusionTechniques

The TBIR and CBIR subsystems are fused in the late multimedia fusion method, whichrefers to the fusion of the TBIR and CBIR subsystems. It offers conclusions that take the shapeof probabilities or ratings that have numerical similarities. Therefore, using the aggregationfunctions Barnard K andForsyth D[8], the probabilities were combined or fused. Algorithmsfor late fusion are superior than those for early fusion.Image re-ranking is a method thatobtains a group of rated items from a textual subsystem and then reorders them in accordancewith the visual score (Pi). The TBIR subsystem's chosen objects are the only ones used by theCBIRsubsystem to computevisual scores (Pi).

## 6.        Dataset

WeutilisedtheCorel-1KandGHIM-10Kdatasets[9,10].Thefirstofthesetwopicturecollections, which was recently utilised, was broken up into 10 separate categories, includingAfricans, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, andFood. One hundred photos with a resolution of 256384 or 384256 pixels make up each class orcategory. The second dataset utilised in this study is GHIM-10K, which covers a variety ofitems and has a total of 10K photos (10 times more than Corel-1K). It is thought to be moredifficult and diversified than Corel-1K. Each group or class has 500 photographs at a resolutionof 300400 pixels or 400300 pixels. There are 10,000 images total, split into 20 groups orcategories.

## 7.        ExpectedResults

It was discovered that in addition to measuring the degree of agreement between several sets ofrankings for the same set of retrieved items, it could also be used to assess the efficacy of variousretrieval procedures. The current investigation utilised the suggested system, where the classes ofimages serve as judges, the average precision objects serving as the present context's judges, andtheaccuracyratesofthethreeseparate approachesservingasthestudy'sobjects.

## 8.        Conclusions

A retrieval technique based on early and late fusion methods was suggested inthiswork. The suggested technique is a straightforward procedure utilised to improve the retrievalperformance of image retrieval based on merging feature descriptors from colour and texturespaces;moreover,itincorporatedcombiningsimilaritymeasurevaluesforthreeknowndistance coefficients. The automated retrieval procedure and the ease of the early and latemerging processes are the suggested method's two primary advantages. It also has a respectableamountofimprovement.Futureresearchshouldbeconductedwithmorepreciseandrepres entative characteristics, and it may also incorporate some additional descriptor featuressuchpictureshapes.

It was discovered that this test could be used to assess the efficacy of various retrievalmethods and that it could also be used to gauge the degree of agreement between several sets ofrankings for the same set of recovered items. In the current investigation, the accuracy rates ofthe three distinct approaches were regarded objects, and the classes of images served as judges,whiletheaverageprecision objects wereseenas judgesin thecurrentcontext.

**References:**

[1] Ali Ahmed, Sara Mohamed ,” Implementation of early and late fusion methods forcontent-basedimageretrieval“,InternationalJournalofAdvancedandAppliedSciences,8(7)2021, Pages: 97-105

[2] Ahmed A (2020). Implementing relevance feedback for content-based medical imageretrieval. IEEE Access, 8: 79969-79976. https://doi.org/10.1109/ACCESS.2020.2990557

[3] Malebary S and Ahmed A (2019). Feature selection and the fusion-based method forenhancing the classification accuracy of SVM for breast cancer detection. InternationalJournalofComputerScienceandNetworkSecurity, 19(11):55-60.

[4] AhmedAandMalebarySJ(2020).Queryexpansionbasedontop-rankedimagesforcontent-basedmedicalimage retrieval.IEEE Access,8:194541-194550.https://doi.org/10.1109/ACCESS.2020.3033504

[5] Latif A, Rasheed A, Sajid U, Ahmed J, Ali N, Ratyal NI, and Khalil T (2019). Content-based image retrieval and feature extraction: A comprehensive review. MathematicalProblemsinEngineering, 2019:9658350.https://doi.org/10.1155/2019/9658350

[6] AlsmadiMK (2020).Content-basedimageretrievalusingcolor,shapeandtexturedescriptors and features. Arabian Journal for Science and Engineering, 45(4): 3317-3330.https://doi.org/10.1007/s13369-020-04384-y

[7] AtreyPK,HossainMA,ElSaddikA,andKankanhalliMS(2010).Multimodalfusionformulti mediaanalysis:Asurvey.MultimediaSystems,16(6):345-379.https://doi.org/10.1007/s00530-010-0182-0

[8] Barnard K and Forsyth D (2001). Learning the semantics of words and pictures. In the8th IEEE International Conference on Computer Vision, IEEE, Vancouver, Canada, 2:408-415.https://doi.org/10.1109/ICCV.2001.937654

[9] BhowmikN,GonzálezR,Gouet-BrunetV,PedriniH,andBlochG(2014).Efficientfusionofmultidimensionaldescriptorsfori mageretrieval.IntheInternationalConferenceonImageProcessing,IEEE,Paris,France:5766-5770.https://doi.org/10.1109/ICIP.2014.7026166

[10] ErkutU,BostancıoğluF,ErtenM,ÖzbayoğluAM,andSolakE(2019).HSVcolorhistogramb asedimageretrievalwithbackgroundelimination.Inthe1stInternationalInformaticsandSoft wareEngineeringConference,IEEE,Ankara,Turkey:1-5.https://doi.org/10.1109/UBMYK48245.2019.8965513