Visual Server

Video and Image Analytics Overview

opentext™

All-In-One Image Processing and Analysis Technology

With the rapid rise of mobile devices, visual content has become extremely valuable for business intelligence. With smart image recognition technology, seemingly low value pictures, such as "selfies," can reveal a great deal about people, their interests, and their environment. While most organizations collect and store images, they do not make use of this valuable resource because of the significant manual efforts required in processing and analysis such data.

OpenText™ Visual Server automates the processing and analysis of image files. It offers a unique combination of a rich set of features, extensive image file type support, and a high level of accuracy in an end-to-end solution. Its capabilities can be grouped into the following categories:

	Optical character recognition	Object detection	Image classification	Object recognition
Description	Extract text from images of machine-printed text	Detect that an object is present	Detect objects of particular categories are in an image or classify objects by feature	Identify a specific object is present
Example	This image has text "The quick brown fox jumps…"	There are three faces in this image.	This image contains a car, a building, and a pedestrian.	The logo shown here is the OpenText logo.

Visual Server offers the following high-level features to help make sense of image data:

Type	Feature	Description
Human	Face detection	Detect faces in an image
	Face recognition	Train and compare faces against a database of known faces
	Face analysis	Analyze faces in images to determine demographic information
	Clothing analysis	Detect clothing and dominant colors of clothing including skin tones

Type	Feature	Description
Object	Object recognition	Detect specific objects such as corporate logos or product packaging
	Object detection	Locate generic objects such as cars, people, chairs, trucks
	Image classification	Categorize different classes of objects depicted in the image
	Change detection	Detect changes in images before and after versions
	Color analysis	Analyze dominant colors of an image
Vehicles	ANPR	Automatically read number plates (license plates) of vehicles
	Vehicle make, model, and color	Recognize the make, model, and color of a vehicle
Text	Optical character recognition (OCR)	Convert text in image files into text files
		Detect text on images such as subtitles on a frame of a video
	1D and 2D barcode detection	Detect barcodes from over 20 barcode types, including ISBN, PDF417, and data matrix
Scenes	Scene analysis	Detect atypical events in surveillance videos such as running, illegal crossings, and traffic violations
	Keyframe analysis	Detect scene transitions in video
Other	Image editing	Blur a region of the image, draw an outline around a region of the image, or crop an image

Supported image types include:

- TIFF
- JPEG
- BMP
- PNG
- GIF
- ICO
- PBM
- PGM
- PPM

Additionally, with the help of OpenText™ KeyView, Visual Server can support document formats including:

- PDF
- DOC and DOCX
- XLS and XLSX
- PPT and PPTX
- ODT
- RTF

Supported video codecs include:

Video Codecs
- MPEG-1
- MPEG-2
- MPEG-4 part 2
- MPEG-4 part 10 (Advanced Video Coding) (H.264)
- MPEG-H part 2 (High Efficiency Video Coding) (H.265)
- Windows media 7
- Windows media 8

File Formats
- MPEG packet stream (for example .mpg)
- MPEG-2 transport stream (for example .ts)
- MPEG-4 (for example .mp4)

- WAVE (Waveform Audio) (.wav)
- ASF (Windows media) (.asf, .wmv)
- Raw AAC (.aac)
- Raw AC3 (.ac3)

Additionally, Visual Server can also ingest video from cameras and third-party video management systems such as:

- MJPEG video streams
- DirectShow device
- Milestone XProtect

Who Benefits?

Many of the Visual Server media analytics features have traditionally been connected to specific industries, for example, facial recognition for security or barcode reading for consumer merchandising. However, as cameras and digital images continue to reach more areas of every-day life, there is the potential for all businesses to leverage the power of image analytics. The ability to run all facets of image analysis together leads to a more efficient and improved workflow.

Features

Visual Server analyzes and edits image files. Visual Server can be used to process large repositories of images and extract information from them. In particular, Visual Server can be used to:

- Detect faces, recognize faces, and analyze faces to extract facial attributes
- Recognize text in scans and photographs of machine-printed text
- Classify images into various object categories
- Locate objects belonging to generic categories within images
- Recognize specific 2D and 3D rigid objects such as movie posters and company logos
- Edit images, crop images, or blur images
- Detect and read barcodes, including QR codes
- Detect most dominant colors present in an image, including skin tones
- Automatically recognize number plates on vehicles
- Recognize vehicle make, vehicle model, and vehicle colors

- Detect atypical events in CCTV camera footage

- Extract keyframes from video

- Generate image hashes to compute approximate image similarity based on color

This section covers the Visual Server features in more detail.

Face Detection

Face detection finds faces in a given image. Visual Server returns the coordinates of a detected face in a photo, as well as the position of key facial features such as the eyes.

Figure 1. Face detection and analysis

Face Recognition

In addition to finding faces, Visual Server and its facial recognition features can also compare the detected face to a database of known individuals. The matching threshold can be adjusted to suit different needs. Visual Server face recognition technology is comparable to other leading face recognition vendors.

Figure 2. Face recognition

Face Analysis

A face can also be analyzed for specific traits. Visual Server can estimate the approximate age range (baby/child/young adult/middle age/elderly), ethnicity, gender, and expression of the person being analyzed.

Object Recognition

Visual Server can be trained to recognize specific objects or complex patterns in analyzed images. For example, a user can train a database of corporate logos to combat copyright infringement. When a picture is analyzed, Visual Server can report if it has found any matching logos from its training set. The objects can be 2D or 3D objects in images.

Figure 3. Logo and object recognition

Image Classification

Image classification automatically categorizes objects that appear in images based on previous training. For example, Visual Server can be trained to recognize vehicle categories such as cars, trucks, and motorcycles. This allows users to sort images as they are analyzed and to flag certain objects, if necessary. Visual Server also provides pre-trained classifiers that can label images with existing categories, so that it becomes easier to automatically tag large collections of images.
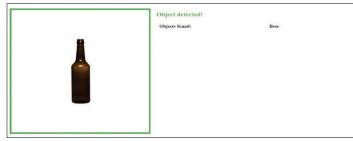

Figure 4. Image classification using trained shapes

Optical Character Recognition (OCR)

Optical character recognition is used to extract text from image files. The use of OCR on scanned or photographed documents, pictures, or photos facilitates the conversion into a computer-readable format to make it easier to store and search the documents. Given an input image, Visual Server can return the identified text, the confidence score, and the region of the image the text was read from. The detection region can be bound to certain areas to decrease noise from the rest of the image and accuracy can be fine-tuned to the position of each character. Visual Server supports all major languages and font types for OCR. A full list of languages and fonts are available in Appendices A and B.

Figure 5. Text is captured from an advertisement using OCR

Barcode

Visual Server has robust support for detecting one-dimensional and two-dimensional barcodes, including QR codes in an image. It can return the data contained in the barcode, as well as barcode type and regions. The following barcode types are supported:

PDF417	Data matrix	ISBN (or EAN-13)	SBN-2 (or EAN-2)	I25
ISBN-5 (or EAN-5)	Code-128	Code-93	Code-39	IATA 2/5
Codabar	Patch Code	Matrix 2/5	Datalogic 2/5	Industrial 2/5
UCC/EAN-128 (or GS1-128)	EAN-8	UPC-A	UPC-E	

Figure 6. Visual Server is able to detect and read barcodes

Object Detection

Object detection can be used to locate instances of objects that belong to known, predefined classes. For example, one could detect all pedestrians, vans, and cars that appear in a video.

Automatic Number Plate Recognition (ANPR)

Detect and read the number plates (license plates) of vehicles in images or video. Number plate recognition has many applications; you can detect stolen and uninsured vehicles, and monitor the length of stay for vehicles in car parks.

Vehicle Make, Model, and Color Analysis

Visual Server can help identify the make, model, and color of a vehicle captured during number plate recognition. Vehicle model recognition can help law enforcement identify stolen vehicles.

Scene Analysis

Scene analysis detects atypical events that occur in video. This can be used to monitor video streamed from CCTV cameras, to assist with the detection of potential threats, illegal actions, or alert human operators to situations where help is required.

Keyframe Analysis

Visual Server can help identify when there are significant scene changes within a video. This can be useful for creating thumbnail photos or time snapshots.

Change Detection

Visual Server can help identify when there are changes within a scene. For example, one may wish to look for objects that have disappeared, new objects that have appeared, or objects that have moved to different parts of an image or scene. This can be used to find defects in images of equipment or create alerts for suspicious movements within a surveillance application.

Color Analysis

Visual Server also includes basic photo analysis functionality, including reporting picture size, color dominancy, and palettes. This is often used in conjunction with object recognition when automating processes that require identification of photo subjects such as cars.

Image Editing

Many analysis tasks return regions of interest. Visual Server provides facility to crop an image to a desired region, blur a region of the image, or draw an outline around a region in the image.

Stand-Alone Server Architecture

Visual Server is a stand-alone media analytics server that uses the OpenText™ Autonomy Content Infrastructure (ACI) Client API to communicate with custom applications. It allows data to be retrieved over HTTP using XML and can adhere to SOAP. It supports both synchronous and asynchronous actions (see the section on scalability).
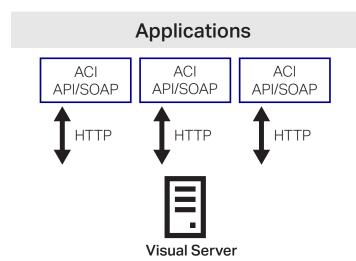

Figure 7. Visual Server architecture

Scalability and High Performance

Visual Server can run several tasks at once in parallel and take full advantage of the available hardware. Tasks can be distributed across several Visual Servers. Visual Server then queues the tasks and runs them in order or multiple tasks at a time. The user can check on the progress of each task, kicking off additional tasks, if necessary, enabling better batch processing and more complicated workflows. Multiple Visual Servers can talk to a common shared database or share data across different databases, so that users get complete flexibility.

Visual Server can accelerate processing by using a GPU. If multiple GPUs are available, one can run multiple Visual Servers on the same machine.

To improve performance in a production environment, Visual Server supports both synchronous and asynchronous actions, and can be distributed for horizontal scaling.

With a synchronous action, Visual Server runs the task immediately and returns a result when the action is complete. Asynchronous actions allow a user to send multiple tasks all at once, returning a task ID/token for each job. In large media analytics systems where a very large number of documents need processing, it is possible to distribute work among multiple instances of Visual Server.

Out-of-the-Box Operation

Several of our analytics come with pre-trained models allowing out-of-the-box operation. We provide pre-trained models for facial demographics, image classification for over 1000 common classes, object detection, pedestrian detection, face detection, vehicle make detection, and vehicle color detection.

Comprehensive Analytics

Visual Server offers full functionality in a single product. A unified solution allows for greater freedom in workflow design and faster integration and deployment. As a result, it is easier to use multiple media analytics using Visual Server and no time is wasted getting multiple vendors' products to work together. A unified solution makes it easy to perform complex analytical queries spanning multiple features.

Summary

In the modern information age, organizations must move beyond merely accessing data to figuring out how to analyze and make sense of vast quantities of data. When businesses can understand information in real

time, it becomes possible to make intelligent, data-driven decisions that can have a positive effect on success rates. Image data is one format that is often stored but not analyzed beyond its simplistic metadata because of the difficulty that traditional technologies have in understanding the vast amount of information held in a single picture. With Visual Server, organizations can improve this increasingly prominent data set to operate at full potential and with greater agility.

Learn more at
www.microfocus.com/richmedia

www.microfocus.com/opentext

Appendix A

Supported Languages for OCR

Latin alphabet	Cyrillic alphabet	Other alphabets
■ Afrikaans (af)	■ Bulgarian (bg)	■ Greek (el)
■ Esperanto (eo)	■ Serbian (sr)	■ Hebrew (he)
■ Italian (it)	■ Macedonian (mk)	■ Arabic (ar)
■ Portuguese (pt)	■ Ukrainian (uk)	■ Persian (fa)
■ Basque (eu)	■ Russian (ru)	■ Urdu (ur)
■ Estonian (et)		■ Japanese (ja)
■ Irish (ga)		■ Simplified Chinese (zhs)
■ Romanian (ro)		■ Traditional Chinese (zht)
■ Catalan (ca)		
■ Finnish (fi)		
■ Latin (la)		
■ Slovak (sk)		
■ Croatian (hr)		
■ French (fr)		
■ Latvian (lv)		
■ Slovenian (sl)		
■ Czech (cs)		
■ German (de)		
■ Lithuanian (lt)		
■ Spanish (es)		
■ Danish (da)		
■ Hungarian (hu)		
■ Maltese (mt)		
■ Swedish (sv)		
■ Dutch (nl)		
■ Icelandic (is)		
■ Norwegian (no)		
■ Turkish (tr)		
■ English (en)		
■ Ido (io)		
■ Polish (pl)		
■ Welsh (cy)		

Appendix B

Supported Special Font and Character Set Codes for OCR

Font
General
Arial Narrow
OCR-A
OCR-B
E13B
Farrington 7B
Old-Style Times
Custom font used for Bloomberg terminal GUI

opentext™