Supporting media bricoleurs with cemint

Media Impact

Where Is the User in Multimedia Retrieval?

Compared to information retrieval or computer vision, multimedia retrieval is a relatively young discipline. Many people mark the 1992 Visual Information Management Workshop at the beginning of the field. It was there that researchers recognized the need to consider multimedia data, in particular visual information, as a new type of item that could appear in a digital collection. Although the number of items at that time was still small, typically in the thousands, it was orders of magnitude larger than the tens of images that computer vision research was addressing. From that first workshop, we have this important point: "Computer vision researchers need to study names require more than just their discipline's current emphasis on automatic techniques, and develop techniques to compare these features in interactive environments." The information retrieval field agreed that new techniques were necessary to cope with the volume of visual data. The notion of visual words, now so popular in visual retrieval, was unheard of at that time. Thus, a new research area was born.

The IBM Query by Image Content (QBIC) system was released not long after the 1992 workshop. QBIC was an early example of a query-by-example system, where the user selected example images or specified the required images. New possibilities for museum collections and medical imaging arose, but techniques were not yet mature enough to have much impact. Various research efforts started to improve humans, especially in terms of their interfaces to various conditions. The early period in content-based retrieval was mainly concerned with understanding the sensor data.

As the mindset that period (around 2000), two new conference series started. The ACM Conference on Multimedia Information Retrieval (ISMIR), which began at the University of Illinois, originally focused on computer vision applications and held its 13th meeting in 2010. The first ACM Conference on Information and Video Retrieval (CVPR) in 2001 had a strong connection to library science, which hosts a community of experts who join data at ingestion time and search for images on request. Until 2007, these conferences always included a few non-technical papers every year that looked at the physical retrieval techniques (such as labeling). The first Video Olympics in 2004, which were interactive systems that announced in front of a live audience of scientists and media librarians, was also held in 2007 at the Sound and Vision Archive in the Netherlands. After 2007, these conferences shifted their focus toward the computational side of the problem, with a stronger emphasis on industrial applications. We can largely attribute this shift to the important role of BIBM. In 2006, the field had a common goal to produce a benchmark to use. Interactive tasks were defined for BIBM, but the concept-detection task especially frustrated. It gave a boost not only in the multimedia retrieval field, but the computer vision community started to introduce the topic as well.

In the early days, we could do what we wanted because we were alone and could easily cater to the task at hand. Then interest came in those days there was no money to buy books. Books you borrowed from the rental library of Shakespeare and Company, which was the library and bookstore of Sylvia Beach at 12 rue de l'Odeon. On a cold wind-swept street, this was a lovely, warm, cheerful place with a big stove in winter, tables and shelves of books, new books in the window, and photographs on the wall of famous writers both dead and living. The photographs all looked like snapshots and even the dead writers looked as though they had really been alive. Sylvia had a lively, very sharply cut face, brown eyes that were as alive as a small animal's and as gay as a young girl's, and wavy brown hair that was brushed back from her fine forehead and cut thick below her ears and at the line of the collar of the brown velvet jacket she wore. She had pretty legs and she was kind, cheerful and interested, and loved to make jokes and gossip. No one that I ever knew was nicer to me.

I was very shy when I first went into the bookshop and I did not have enough money on me to join the rental library. She told me I could pay the deposit any time I had the money and made me out a card and said I could take as many books as I wished.

There was no reason for her to trust me. She did not know me and the address I had given her, 74 rue Cardinal Lemoine, could not have been a poorer one. But she was delightful and charming and welcoming and behind her, as

Editor's Note
This article summarizes a round-table discussion at the ACM International Conference on Multimedia Retrieval, where a case was made for making the interacting user a first-class citizen again in multimedia retrieval research.
6 billion hours of video are watched each month on YouTube ... 100 hours of video are uploaded to YouTube every minute
Best possible way to communicate ideas is to be able to use a combination of media

<table>
<thead>
<tr>
<th>Content</th>
<th>Media</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedures</td>
<td>Animation</td>
</tr>
<tr>
<td>Analysis</td>
<td>Text</td>
</tr>
<tr>
<td>Representation</td>
<td>Photos</td>
</tr>
</tbody>
</table>
Why?

Suni Brown argues that doodle or sketch is "to make spontaneous marks to help yourself think." SketchScan augments this process, helping others (as well as your future self) see how you came up with ideas through the marks you made.

Other ways you can use SketchScan

- Annotate non-sketched images
- Use text-to-speech to create videos anytime
- Or just use it to clean and archive photos of sketches

How?

1. Capture

The first step is to capture an image.
<table>
<thead>
<tr>
<th>Effect</th>
<th>User driven task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Temporal</strong></td>
</tr>
<tr>
<td><strong>Temporal</strong></td>
<td>Play animation at video time</td>
</tr>
<tr>
<td><strong>Spatial</strong></td>
<td>Advance slide at video time</td>
</tr>
</tbody>
</table>
“Come a week early,” my pilot said, “so you can learn how to fly the airplane.”

Chapter 1
City of Dogs
Wolf Lake

The Last Great Race
The snow burst through the trees with no warning but a last-second whoosh of sound, a two-story wall of white and Chris Rudolph’s piercing cry: “Avalanche! Eyewot!”

The only thing the 16 skiers and snowboarders had sought — fresh, soft snow — instantly became the enemy. Somewhere above, a pristine meadow cracked in the shape of a lightning bolt, slicing a slab nearly 200 feet across and 3 feet deep. Gravity did the rest.

Snow shattered and spilled down the slope. Within seconds, the avalanche was the size of more than a thousand cars barreling down the mountain and weighed millions of pounds. Moving about 70 miles per hour, it crashed through the sturdy old-growth trees, snapping their limbs and shredding back from their trunks.

The avalanche, in Washington’s Cascades in February, did not some trees and rocks, like ocean swells around a ship’s prow. Others it captured and added to its violent load.

Somewhere inside, it also carried people. How many, no one knew.

The slope of the terrain, shaped like a funnel, squeezed the growing swell of churning snow into a steep, twisting gorge. It moved in surges, like a roller coaster on a series of drops and high-banked turns. It accelerated as the slope steepened and the weight of the slide pushed from behind. It flattened through shallower pitches. The energy raised the temperature of the snow a couple of degrees, and the friction carved striations high in the icy sides of the canyon walls.

Elise Saugstad, a professional skier, wore a backpack equipped with an air bag, a relatively new and expensive part of the arsenal that backcountry users increasingly carry to ease their minds and increase survival odds in case of an avalanche. About to be overtaken, she pulled a cord near her chest. She was knocked down before she knew if the canister of compressed air inflated her inflated pillows behind her head.

She had no control of her body as she tumbled downhill. She did not know up from down. It was not unlike being earth-headed in a relentlessly crashing wave. But snow does not need. It swallows its victims. It does not spit them out.

Snow filled her mouth. She screamed off things she never saw, tumbling through a shattered canyon like a steel marble falling through pins in a pachinko machine.

At first she thought she would be embarrassed that she had deployed her air bag, that the other expert skiers she was with, more than a dozen of them, would have a good laugh at her panicked overreaction. Seconds later, tumbling uncontrollably inside a ribbon of speeding snow, she was sure this was how she was going to die.

Moving, rolling snow turns into something closer to liquid, thick like lava. But when it stops, it instantly freezes solid. The laws of physics and chemistry transform a meadow of fine powder into a wreckage of icy chunks. Saugstad’s pinwheeling body would freeze into whatever position it was in the moment the snow stopped.
<table>
<thead>
<tr>
<th>Effect</th>
<th>Cause</th>
<th>Temporal</th>
<th>Spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal</td>
<td>Play animation at video time</td>
<td>Play animation when navigate to part of article</td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Advance slide at video time</td>
<td>Update map when navigate to part of article</td>
<td></td>
</tr>
</tbody>
</table>

Need better creation tools!
“... in the most fundamental sense, we, as learners, are all bricoleurs.”

Build our understanding of complicated processes by reconfiguration, tinkering

To tinker you need building blocks and components

Word processors great for this, not such support for multimedia
Carbon dioxide is converted into sugars in a process called [[carbon fixation]]. Carbon fixation is an [[endothermic]] [[redox]] reaction, so photosynthesis needs to supply both a source of energy to drive this process, and the electrons needed to convert carbon dioxide into a [[carbohydrate]]. This addition of the electrons is a [[Reduction (chemistry)|reduction reaction]]. In general outline and in effect, photosynthesis is the opposite of [[cellular respiration]], in which glucose and other compounds are oxidized to produce carbon dioxide and water, and to release [[exothermic]] chemical energy to drive the organism's [[metabolism]]. However, the two processes take place through a different sequence of chemical reactions and in different cellular compartments.

The general [[chemical equation|equation]] for photosynthesis is therefore:
\[2n \text{ CO}_2(s) + 2n \text{ H}_2\text{O}(l) + \text{[photons]} \rightarrow \text{[carbohydrate]}_2(\text{CH}_2\text{O})_n(s) + \text{2n DO}\]

Carbon dioxide + electron donor + light energy → carbohydrate + oxidized electron donor

In "oxygenic" photosynthesis water is the electron donor and, since its [[hydrolysis]] releases oxygen, the equation for this process is:
\[2n \text{ CO}_2(s) + 4n \text{ H}_2\text{O}(l) + \text{[photons]} \rightarrow \text{[carbohydrate]}_2(\text{CH}_2\text{O})_n(s) + \text{2n O}_2(s)\]

Carbon dioxide + water + light energy → carbohydrate + oxygen + water

Often 2n water molecules are cancelled on both sides, yielding:
\[2n \text{ CO}_2(s) + 2n \text{ H}_2\text{O}(l) + [\text{photons}] \rightarrow \text{[carbohydrate]}_2(\text{CH}_2\text{O})_n(s) + \text{2n O}_2(s)\]

Carbon dioxide + water + light energy → carbohydrate + oxygen
“In my initial requirements for this thing, I had assumed, as an absolute pre-condition, that nobody would have to do HTML or deal with URLs. If you use the original World Wide Web program, you never see a URL or have to deal with HTML. You're presented with the raw information. You then input more information. So you are linking information to information--like using a word processor.”

Tim Berners-Lee
component extraction from media for interaction, navigation, and transformation
Content-based Copy and Paste from Video Documents

```javascript
// TODO: keep highlighted boxes up to date based on mousedown and mousemove
function canvasCapture()
{
    var overlay = null;
    var FONT_SIZE = "14px Arial";
    var SKIP_FRAMES = 1;
    var async = false;
    var DETECT_SCROLL = true;
    var DETECT_FACES = false;
    var JSade = null;
    var enableBoxSelection = false;
    var videoEl = "video";
    var scaleFactor = 0.56/0.33/0.56/0.25;
    var snapshots = [];
    var canvas = document.createElement("canvas");
    var fakecanvas = null;
    var processedcanvas = null;
    var processedcanvas2 = null;
    var processedcanvas3 = null;
    var annotations = [];
    var width = -1;
    var height = -1;
    var video = null;
}
```
User double taps the dialog box and...
User double taps the dialog box and... ...the video canvas zooms into the dialog box
We are in a video conference with Sharon
cemint
cemint
Use WebRTC + real time image analysis to determine important points for navigation and control.
Summary

Best possible way to communicate ideas is to be able to use a combination of media

We don’t really have all the tools to facilitate this yet

cemint is one start
“If your medium doesn't easily allow you to correct mistakes, knowledge will tend to be carefully vetted. If it's expensive to publish, then you will create mechanisms that winnow out contenders. If you're publishing on paper, you will create centralized locations where you amass books... Traditional knowledge has been an accident of paper.”

David Weinberger
“If your medium doesn't easily allow you to correct mistakes, knowledge will tend to be carefully vetted. If it's expensive to publish, then you will create mechanisms that winnow out contenders. If you're publishing on paper, you will create centralized locations where you amass books... Traditional knowledge has been an accident of paper.”

David Weinberger

Questions?
User mouse-wheels up over iPhone screen
User mouse-wheels up over iPhone screen

Video automatically seeks backward
Step 1. Check that printer is off

Step 2. Locate outer panel access button