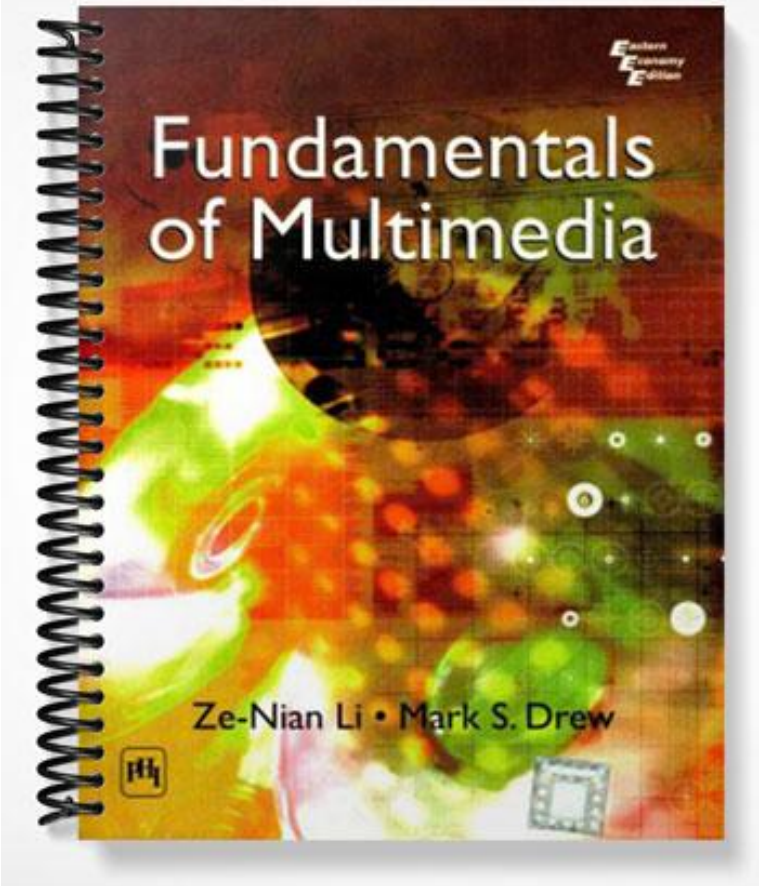
SOLUTIONS MANUAL



Fundamentals of Multimedia
ISBN: 0130618721

Ze-Nian Li and Mark S. Drew
School of Computing Science
Simon Fraser University

Exercise Solutions

Contents

Chapter 1

Introduction to Multimedia

Exercises

1. Identify three novel applications of the Internet or multimedia applications. Discuss why you think these are novel.

   Answer:
   WML – Wireless markup Language.
   Mobile games: massive multiplayer online role-playing game (Mmorpg)
   Multisensory data capture
   Capture of context
   Represent and adjust recollections in memory over time
   "Fictive Art" in new media: beyond RPGs in the use of narrative and fictions to create made-up worlds, imaginary situations, and odd situations (an example is "The Museum of Jurassic Technology").
   Bridging the semantic gap problem in automatic content annotation systems — the gulf between the semantics that users expect and the low-level features (content descriptions) that systems actually use: one solution is "an approach called computational media aesthetics. We define this approach as the algorithmic study of a variety of image, space, and aural elements employed in media ... based on the media elements usage patterns in production and the associated computational analysis of the principles that have emerged while clarifying, intensifying, and interpreting some event for the audience." (IEEE Multimedia Magazine, Volume: 10, Issue: 2, Year: April-June 2003)

2. Briefly explain, in your own words, "Memex" and its role regarding hypertext. Could we carry out the Memex task today? How do you use Memex ideas in your own work?

   Answer:
   Memex was a theoretical system explicated by Vanevar Bush in a famous 1945 essay. His main ideas involved using *associative memory* as an aid for organizing a weler of material. He even adumbrated the concept of *links*.

3. Your task is to think about the transmission of smell over the Internet. Suppose we have a smell sensor at one location and wish to transmit the *Aroma Vector* (say) to a receiver to reproduce the same sensation. You are asked to design such a system. List three key issues to consider and two applications of such a delivery system. *Hint*: Think about medical applications.

Answer:
"October 6, 2000 – DigiScents, Inc., the pioneer of digital scent technology, received the 'Best New Technology' award for its iSmell(TM) device at the 'Best of RetailVision Awards' at the Walt Disney World Dolphin Hotel in Lake Buena Vista, Fla. Retailers such as BestBuy, RadioShack, CompUSA and other industry giants voted on the vendor awards."
"DigiScents ... The company would send you a dispenser about the size of a computer speaker. You'd plug it into your PC. It would be filled with chemicals that, when mixed, could recreate most any smell. Tiny bits of data would come in over the Net to tell your dispenser what smell to make. There would be a portal where you could find scents. DigiScents calls it – and at first I thought they were joking – a 'Snortal.'"

4. Tracking objects or people can be done by both sight and sound. While vision systems are precise, they are relatively expensive; on the other hand, a pair of microphones can detect a person's *bearing* inaccurately but cheaply. Sensor *fusion* of sound and vision is thus useful. Surf the web to find out who is developing tools for video conferencing using this kind of multimedia idea.

   Answer:
   "Distributed Meetings: A Meeting Capture and Broadcasting System," Ross Cutler, Yong Rui, Anoop Gupta, JJ Cadiz Ivan Tashev, Li-wei He, Alex Colburn, Zhengyou Zhang, Zicheng Liu, Steve Silverberg, Microsoft Research, ACM Multimedia 2002,
   http://research.microsoft.com/research/coet/V-Kitchen/chi2001/paper.pdf

5. *Non-photorealistic* graphics means computer graphics that do well enough without attempting to make images that look like camera images. An example is conferencing (let's look at this cutting-edge application again). For example, if we track lip movements, we can generate the right animation to fit our face. If we don't much like our own face, we can substitute another one — facial-feature modeling can map correct lip movements onto another model. See if you can find out who is carrying out research on generating avatars to represent conference participants' bodies.

   Answer:
   See: anthropic.co.uk

6. Watermarking is a means of embedding a hidden message in data. This could have important legal implications: Is this image copied? Is this image doctored? Who took it? Where? Think of "messages" that could be sensed while capturing an image and secretly embedded in the image, so as to answer these questions. (A similar question derives from the use of cell phones. What could we use to determine who is putting this phone to use, and where, and when? This could eliminate the need for passwords.)

   Answer:
   Embed retinal scan plus date/time, plus GPS data; sense fingerprint.

Chapter 2

Multimedia Authoring and Tools

Exercises

1. What extra information is multimedia good at conveying?

   (a) What can spoken text convey that written text cannot?

   Answer:
   Speed, rhythm, pitch, pauses, etc...
   Emotion, feeling, attitude ...

   (b) When might written text be better than spoken text?

   Answer:
   Random access, user-controlled pace of access (i.e. reading vs. listening)
   Visual aspects of presentation (headings, indents, fonts, etc. can convey information)
   For example: the following two pieces of text may sound the same when spoken:
   I said "quickly, come here."
   I said quickly "come here."

2. Find and learn 3D Studio Max in your local lab software. Read the online tutorials to see this software's approach to a 3D modeling technique. Learn texture mapping and animation using this product. Make a 3D model after carrying out these steps.

3. Design an interactive web page using Dreamweaver. HTML 4 provides layer functionality, as in Adobe Photoshop. Each layer represents an HTML object, such as text, an image, or a simple HTML page. In Dreamweaver, each layer has a marker associated with it. Therefore, highlighting the layer marker selects the entire layer, to which you can apply any desired effect. As in Flash, you can add buttons and behaviors for navigation and control. You can create animations using the Timeline behavior.

4. In regard to automatic authoring,

   (a) What would you suppose is meant by the term "active images"?

   Answer:
   Simple approach: Parts of the image are clickable.
   More complex: Parts of the image have knowledge about themselves.

   (b) What are the problems associated with moving text-based techniques to the realm of image-based automatic authoring?

Answer:
Automatic layout is well established, as is capture of low-level structures such as images and video. However amalgamating these into higher-level representations is not well understood, nor is automatically forming and linking appropriate level anchors and links.

(c) What is the single most important problem associated with automatic authoring using legacy (already written) text documents?

Answer:
Overwhelming number of nodes created, and how to manage and maintain these.

5. Suppose we wish to create a simple animation, as in Fig. 2.30. Note that this image is exactly

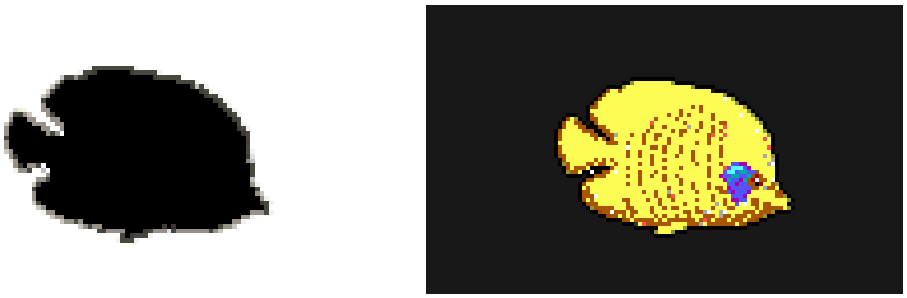
Fig. 2.30: Sprite, progressively taking up more space.

what the animation looks like at some time, not a figurative representation of the *process* of moving the fish; the fish is repeated as it moves. State what we need to carry out this objective, and give a simple pseudocode solution for the problem. Assume we already have a list of (x, y) coordinates for the fish path, that we have available a procedure for centering images on path positions, and that the movement takes place on top of a video.

Answer:

```
\\ We have a fish mask as in Figure \ref{FIG:MASKANDSPRITE}(a), and
\\ also a fish sprite as in Figure \ref{FIG:MASKANDSPRITE}(b).
\\ Fish positions have centers posn(t).x posn(t).y

currentmask = an all-white image
currentsprite = an all-black image
for t = 1 to maxtime {
   \\ Make a mask fishmask with the fish mask black area
   \\ centered on position posn(t).x, posn(t).y
   \\ and a sprite fishsprite with the colored area also moved
   \\ to posn(t).x, posn(t).y
   \\ Then expand the mask:
   currentmask = currentmask AND fishmask \\ enlarges mask
   currentsprite = currentsprite OR fishsprite \\ enlarges sprite
   \\ display current frame of video with fish path on top:
```

```
    currentframe = (frame(t) AND currentmask) OR currentsprite
}
```



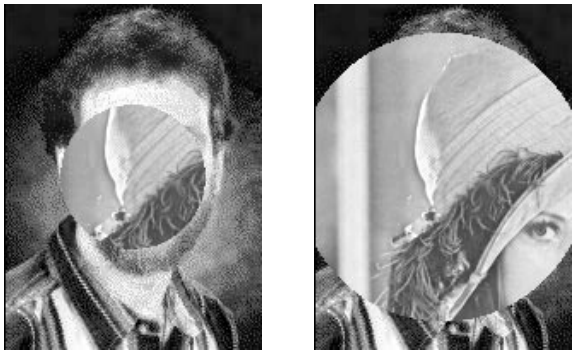
(a) (b)

Fig. 2.30: (answer) Mask and Sprite.

6. For the slide transition in Fig. 2.11, explain how we arrive at the formula for x in the unmoving right video R_R.

   Answer:
   if $x/x_{max} \geq t/t_{max}$, then we are in the right-hand video. The value of x is to the right of x_T, and the value in the *unmoving* right-hand video is that value of x, reduced by x_T so that we are in units with respect to the left of the right-hand video frame. That is, in the right-hand video frame we are at position $x - x_t$, which is $x - (x_{max} * t/t_{max})$.

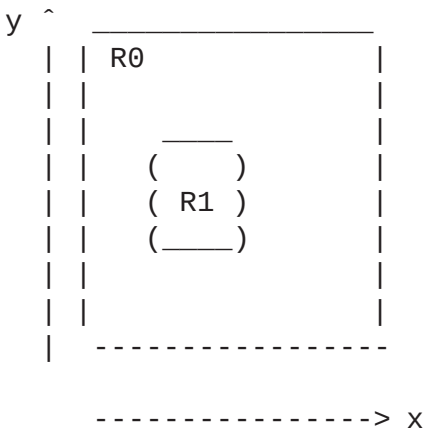
7. Suppose we wish to create a video transition such that the second video appears under the first video through an opening circle (like a camera iris opening), as in Figure 2.31. Write a formula to use the correct pixels from the two videos to achieve this special effect. Just write your answer for the red channel.



(a) (b)

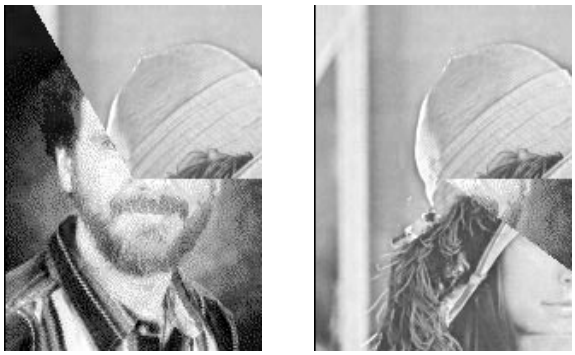
Fig. 2.31: Iris wipe: (a): Iris is opening. (b): At a later moment.

Answer:




```
radius of transition r_T = 0 at time t = 0

r_T = r_max = sqrt( (x_max/2)^ + (y_max/2)^2 )  at time t=t_max

-->  r_T = r_max * t / t_max

At x,y,
  r = sqrt(  (x-x_max/2)^2 + (y-y_max/2)^2  )

  If (  r < (t/t_max)*r_max  )
     R(x,y,t) = R1(x,y,t)
  Else
     R(x,y,t) = R0(x,y,t)
```

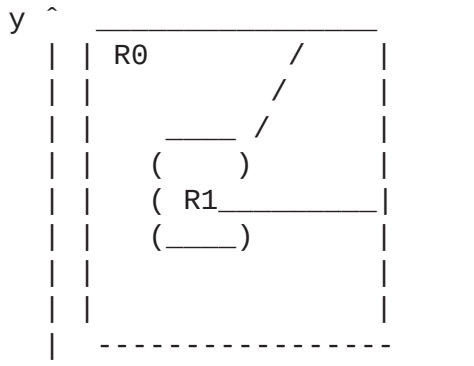
8. Now suppose we wish to create a video transition such that the second video appears under the first video through a moving radius (like a clock hand), as in Figure 2.32. Write a formula to use the correct pixels from the two videos to achieve this special effect for the red channel.



(a) (b)

Fig. 2.32: Clock wipe: (a): Clock hand is sweeping out. (b): At a later moment.

Answer:




```
angle of transition a_T = 0 at time t = 0

a_T = a_max = 360  at time t=t_max

-->  a_T = a_max * t / t_max


At x,y,
  a = atan(  -(y-y_max/2)/(x-x_max/2)  )
  \\ Since y in defn of angle increases from bottom, not from top
  \\ like rows. Watch for correct quadrant, though--use 2-arg atan.

  If (  a < (t/t_max)*a_max  )
     R(x,y,t) = R1(x,y,t)
  Else
     R(x,y,t) = R0(x,y,t)
```

9. Suppose you wish to create a wavy effect, as in Figure 2.33. This effect comes from replacing the image x value by an x value offset by a small amount. Suppose the image size is 160 rows $\times 120$ columns of pixels.

   (a) Using float arithmetic, add a sine component to the x value of the pixel such that the pixel takes on an RGB value equal to that of a different pixel in the original image. Make the maximum shift in x equal to 16 pixels.

   Answer:
   R = R(x + sin(y/120) * 16 , y) and similarly for G, B.

   (b) In Premiere and other packages, only integer arithmetic is provided. Functions such as `sin` are redefined so as to take an `int` argument and return an `int`. The argument to the `sin` function must be in $0\,..\,1{,}024$, and the value of `sin` is in $-512\,..\,512$: `sin(0)` returns `0`, `sin(256)` returns `512`, `sin(512)` returns `0`, `sin(768)` returns `-512` and `sin(1,024)` returns `0`. Rewrite your expression in part (a) using integer arithmetic.

Fig. 2.33: Filter applied to video.

Answer:

```
R = R(x + sin( (y*1024)/120 ) /32,y) and similarly
for G,B.
[In Premiere:     src(x + sin( (y*1024)/120 ) /32,y ,p)    ]
Why: y in 0..119;  (y*1024)/120 in 0..1023; the resulting  sin
is in -512..512;  and dividing by 32 puts the offset range into
-16..16.
```

(c) How could you change your answer to make the waving time-dependent?

Answer:

```
R = R(x +  t*sin(y*(1024/120) ) /(32*tmax),y)
[In Premiere: src(x +  t*sin(y*(1024/120) ) /(32*tmax),y ,p) ]
Note the order: else have t/tmax==0 or 1 only.
```

10. How would you create the image in Figure 2.6? Write a small program to make such an image. *Hint:* Place R, G, and B at the corners of an equilateral triangle inside the circle. It's best to go over all columns and rows in the output image rather than simply going around the disk and trying to map results back to (x, y) pixel positions.

Answer:

```
% Matlab script:
SIZE = 256;
im = ones(SIZE,SIZE,3);
% Place R at (0,1).
```

```
% Place G at 120 degrees.
% Place B at 240 degrees.
% The outside perimeter goes from R to G as we go from
%   R to G.
% And from B to R as we go from 240 to 360.
%
% At a position where the Outside Perimeter value
% is   out  , at radius   r  the color is
% (1-r)*(1,1,1) + r*(out)

% Go over all pixels:
for j = 1:SIZE
for i = 1:SIZE
  x = j-SIZE/2;
  y = i-SIZE/2;
  r = sqrt(x*x+y*y);
  if (r<=(SIZE/2))
    ang = 180/pi*atan2(y,x);
    if ang <0
      ang = 360+ang;
    end
    if ang < 120 % between R and G
      out = [(120-ang)/120 ; ang/120 ; 0];
    elseif ang < 240 % between G and B
      out = [0 ; (240-ang)/120 ; (ang-120)/120];
    else % between B and R
      out = [(ang-240)/120 ; 0 ; (360-ang)/120];
    end; % if ang
    % and could make the in-between bands broader by not using
    % linear interpolation, if wished.

    %linear:
      im(i,j,:) = ((SIZE/2)-r)/(SIZE/2)*[1;1;1] + r/(SIZE/2)*out;
    % and normalize the color to bright:
    temp = max( im(i,j,:) );
    im(i,j,:) = im(i,j,:)/temp; % takes one channel to 1.0
  end; % if r

end
end
imshow(im)

imwrite(im,'colorwheel256.bmp');
```

11. As a longer exercise for learning existing software for manipulating images, video, and music, make a 1-minute digital video. By the end of this exercise, you should be familiar with PC-based equipment and know how to use Adobe Premiere, Photoshop, Cakewalk Pro Audio, and other multimedia software.