Video: Measuring Image Motion

[00:01] [slide 1] Our next topic in the course is the analysis of visual motion. This first video provides a brief overview of motion analysis and describes the motion information that we can measure directly from the changing image. The second video gets more deeply into the computation of a velocity field that captures the movement of features in the image. Along the way we'll highlight an aspect of motion processing that may underlie some perceptual illusions, and also touch on the neural processing of motion in an area of the primate brain known as MT.

[00:38] The analysis of motion can be divided into two main stages - the first is to determine how features are moving in the two-dimensional image and the second is to use the measurements of image motion to determine the three-dimensional structure of the scene, how objects are moving in space, and how we're moving relative to the environment. The picture in the top left corner here illustrates the kind of information that we might derive from the first stage of motion processing. It's a snapshot from a video taken from an airplane that's banking to the right at this moment in time, and the arrows superimposed on the picture indicate the direction and speed of motion of different parts of the image as the plane flies along. We refer to this as a velocity field. The cylinder in the middle also has arrows superimposed that show how the surface texture moves in the image as the cylinder rotates in space around a central vertical axis.

[01:48] We use these measurements of image motion to perform tasks like detecting object boundaries from the changes in velocity that occur around the borders of objects that are moving relative to a background, like the eyes and nose of the happy face here. In the case of the rotating cylinder, the speed of image movement is conveyed by the length of the arrows, and if you look closely, you can see that the speed varies across the horizontal extent of the object, and we use these variations in image movement to infer the three-dimensional shape of objects like this from their two-dimensional projection onto the image. Analysis of the subtle 3D movements of surfaces is also important for tasks like understanding facial expressions. In the bottom left corner here, as Luke & Leia are zipping through the forest on their speed bikes, they're able to use the expanding pattern of image motion to determine where they're heading in the scene, so they can make quick adjustments to avoid objects like this big tree. We can also infer how other objects are moving in space relative to us, and we're especially sensitive to looming motion when an object is heading directly toward us like this baseball. We're far from perfect at measuring image motion and you'll learn about some motion illusions that arise when we don't quite get it right.

[03:30] [slide 2] When we observe a dynamically changing image, either viewing motion in the real world or in a movie, we can't directly measure the true motion of image features by analyzing what's happening in small regions of the image, because of a problem known as the aperture problem. To get a sense of what I mean, we'll look at a very simple demonstration in MATLAB. Here there's just a line and we're viewing through an aperture, and let's set it in motion. You're probably thinking it's just moving back and forth in an oblique direction, perpendicular to its orientation. But let's observe the motion again with no aperture. As you can see, it's really moving back and forth in the horizontal direction, but we can't figure that out just from information we observe through the aperture. There's actually a continuum of different directions that the line could be moving, but you can't distinguish between them from the information you observe through this aperture. We could add texture to the line, like a dot, for example, and if we interpret this dot as being fixed to the line, you can now perceive the correct horizontal motion back and forth, but all we can really sense directly, if we go back to our original situation, all we can sense directly is its movement in the direction perpendicular to its orientation. This isn't a limitation that just exists in this toy world - wherever you have an extended edge in the image, whose motion is analyzed by observing changes taking place in a small region of the image, you face this limitation. And in most computer vision applications, motion is initially detected by analyzing small regions of the image, as we did in the case of edge detection and stereo processing.

[05:55] There's another complicating factor with motion - it's impossible, in general, to identify a unique pattern of motion from the changing image. If features in the image change their appearance over time, as in the case of the changing contour on the left, a particular point at time 1 could have moved to any point on the contour at time 2. Yet when we observe this contour in motion, we'll see it moving a particular way, and we'll probably all see it moving the same way. In order to assign a particular direction and speed of movement to each point on the contour, we need additional constraints that enable us to compute a unique pattern of motion that makes sense from a physical standpoint. This is much like assuming uniqueness or continuity in the case of stereo matching.

[06:58] [slide 3] One common assumption that's made in motion analysis is that all the features within a small region of the image are moving in the same way - we assume that objects are undergoing pure translation over time, and not rotating or distorting, for example. How can this assumption help us to compute the motion of a figure like this square, if we can only directly measure how the edges are moving in the direction perpendicular to their orientation. (And we'll ignore the corners for now.) Let's say we measure the perpendicular motion for the edge circled in green, and that it has this speed shown with the blue arrow. There's a family of possible velocities that this could be consistent with - they're all the motions that we can draw from the point on the edge to the dashed green line here. Every one of these velocities, for example, has a component in the direction perpendicular to the edge that is this blue vector that we measured. Now consider the other edge, circled in red, and suppose that we determined that the motion perpendicular to that edge has this speed shown with the new blue arrow. From this observation, we can say that there's a family of possible velocities for the object that are all the motions that we can draw from the point on this edge to the dashed red line - these vectors, for example, all have the same component in this perpendicular direction.

[08:49] If we consider these two components together, can we resolve the true motion of the object, if it's really just translating across the image? The answer is yes, and to see this, we'll bring those two constraints to a common origin in a coordinate frame that we'll call velocity space. Here, the x and y axes correspond to the horizontal and vertical components of velocity. We'll first draw the constraint we have from the edge with the green aperture, and I drew it slightly enlarged here. In this space, the possible velocities that are consistent with this measurement are vectors that we can draw from the origin to a point on this green line. We can then add the constraint we have from the edge with the red aperture, which says that the possible velocities are vectors that we can draw from the origin to a point on this red line. Is there a velocity that's consistent with both measurements? Fortunately yes, it's the velocity vector from the origin to the point where the two constraint lines intersect - the horizontal velocity that's the true movement of the square in this case, and it's shown with the bold black arrow on the horizontal axis. This construction is sometimes referred to as the "intersection of constraints" for the solution to the aperture problem.

[10:24] [slide 4] You may have a little sense of deja vu about the assumption we're using here - you've used an assumption like this in other contexts. In the fingerprint matching problem, you had a partial print that was a tiny patch of a mystery fingerprint, and you searched around a large fingerprint image to find a patch that closely resembles your partial print. In the case of stereo matching, you explored a region-based method that considered small patches of the left image and searched along a horizontal line in the right image for a matching patch. In the stereo context, you're assuming that the patch looks the same in the right image, and it's just a uniform shift relative to its appearance in the left image. In the fingerprint case, you're assuming that the partial print will look exactly the same in the full fingerprint on file, there won't be any rotation or smudging or distortion of this part of the print. Maybe there's lightbulbs going off in your head - you could implement a motion measurement strategy where you have two images now separated in time, and for each patch in one image, you look for a similar patch in the next, assuming that the patch just translated to a new location. The problem is complicated by the fact that now, the patch could move with any direction and speed in two dimensions, so you'll need to broaden the region where you search for a match, to a two-dimensional area of the next image in the sequence. This is similar to what you did for fingerprint matching. But keep in mind that you'll still need to deal with the aperture problem. These stereo images here are pretty textured, but wherever you have straight edges that extend beyond the patch size, you can't resolve the real motion of features through measurements over small areas of the image.

[12:37] [slide 5] There are a couple other practical considerations that arise. There's likely to be error in the initial motion measurements - errors in the measurement of the perpendicular components of motion, or fluctuations of brightness within patches of the image that may be due to factors like noise in the sensors. The next point is just another way to express the aperture problem - it arises when there's very little variation in the orientation of image texture within a region of the image. The last point is more critical - velocities may not be constant locally, as assumed. If an object is rotating in space, or it changes shape over time, the motion in the image can be different from one location to the next, and we take advantage of these differences, for example, to infer the three-dimensional shape of the object, or to recognize a facial expression, so it's important to capture these subtle variations, and they could be lost if we assume extended regions of the image undergo a uniform motion. But, strategies that assume pure translation can still be useful for tasks like detecting sudden movements, tracking objects through a scene, and detecting object boundaries from motion changes.

[14:10] [slide 6] For the rest of this video, our goals are to learn how we can measure the components of motion in the direction perpendicular to moving edges in the image, and how we can formally express the constraint that these measurements impose on the possible velocities we could assign to a region of the image. You'll also see how we can compute this intersection of constraints from two measurements of these components, but this isn't how we're really going to apply these ideas in practice - for that, I'll leave you hanging until the second video.

[14:50] [slide 7] We'll start by looking at how we can measure motion in one dimension. In order to detect movement at all, the intensity needs to be changing in the image and over time - if we're looking at a large white wall, and it moves, we'd never know it, because nothing changes. So we know there needs to be a change of intensity in the image in order to detect movement, so let's start there. Here's a simple intensity profile where intensity is increasing on the left as we move across the image, and decreasing on the right, and imagine that the intensity on the vertical axis goes from black to white. Now consider what happens when the pattern moves. We'll let Vx refer to its velocity in the x direction. If it moves to the right, we'll say that Vx > 0, and if it moves to the left, we'll say that Vx < 0. Now think about what happens over time, at a particular location in the image, when the pattern moves. We'll first look at a location in the middle of the intensity change on the left, that's marked with the blue bar here. What happens to the intensity at this particular location, when the pattern moves to the right, as shown with the green curve? Before it started to move, the brightness at that point was a medium gray here, but as it moved, brightness decreased at that location, down to a dark gray at this moment in time. Suppose the pattern instead shifted to the left, as shown with the red curve here. At this same location, it started out as the medium gray, but then we would see intensity rise, from the medium gray up to this light gray at this particular moment in time here.

[17:08] I'm going to record what we observe in this table here. The quantity at the top is the derivative of the intensity with respect to x - how it's changing in the image at this location that we're observing here. In the left column of the table we'll record what happens when this derivative is positive, like it is on the left side of the pattern. In the right column we'll record what happens when this derivative is negative, as it is on the right side of the pattern. The label on the left side of the table is the derivative with respect to time - how did the intensity at a particular location like this blue bar here change over time when the pattern moved? In the top row, we'll record situations where intensity increased over time, and in the bottom row, we'll record situations where intensity decreased over time. So what we said so far is that if intensity is increasing in the image, which is what we have on the left side here, then if it decreased over time then the pattern was moving to the right, and we'll record that by putting a right arrow in that spot of the table because intensity change in the image is positive and it decreased over time. On the other hand, in the situation where intensity increased over time, the pattern was moving to the left, so we'll put a red arrow in that spot of the table, a red arrow pointing to the left.

[19:02] Now what happens at a particular location on the other side of the pattern, where intensity is decreasing in the image and this derivative with respect to x is negative, so we're in the right column here. When the profile shifts to the right, again shown with the green curve, the intensity at this location here increases from this medium gray up to this light gray up here, so we're going to put a right moving arrow in that spot in our table. Finally, if it's shifting to the left, at that particular location, intensity will decrease, so we'll note that with a leftward moving red arrow here. So far, we can say that by looking at the sign of how intensity is changing in the image and the sign of how it's changing over time, at a particular location, we can infer the direction of motion of the intensity pattern. But we can be more quantitative here. We can infer both the direction and speed of motion, where we're expressing speed in units of pixels per time step. The velocity of the pattern in the horizontal direction is the rate of change of intensity over time divided by the rate of change in the image. The rate of change in the image is the slope of the intensity profile, and the rate of change over time is the change of intensity at a single location, over time. And there's a negative sign here - we saw, for example, that when both quantities are positive, the pattern is moving to the left, which is a negative velocity.

[21:02] For a bit of intuition here - at the locations of the blue bars, if intensity is changing very slowly over time, the edge must be moving very slowly. In line with that intuition, the time derivative in the numerator would be small in this case, so the computed velocity would also be small. If the intensity is changing very rapidly, the edge must be whizzing by this location very quickly. In this case, the numerator is large, so the computed velocity is large as well.

[21:40] [slide 8] Now, how do we perform these computations in two dimensions? We'll elaborate on these computations through a simple two-dimensional extension of our one-dimensional example. Now imagine that we have extended edges in two dimensions, and this tilted image patch here has a cross-section that's the same intensity profile that we saw in the one-dimensional case, shown on the right. We'll assume a coordinate frame with x increasing to the right and y increasing downward, and let's again consider what happens at a particular location in the image when the pattern moves, so we'll consider the location that's indicated by this red dot. Imagine that the true direction and speed of motion of this edge is given by the green vector here - the patch is really moving horizontally. We know that it's only possible to directly measure the component of this velocity in the direction perpendicular to the edge, which is in the direction of this blue vector here, the motion component. How do we represent this information, and how do we compute it from the changing image? We'll apply the same general concepts that we introduced in one dimension. We'll measure how the intensity pattern is changing in the image and over time, and combine those two quantities to determine the direction and speed of motion, now it will be along this direction perpendicular to the edge. First we need a way to describe the direction that intensity is changing in the two-dimensional image, and for this, we'll use the gradient of intensity. The gradient is a vector that points in the direction of steepest increase of a function. I'm going to add two hot pink vectors, just for a moment, to illustrate the direction of the gradient at two locations. The fact that the gradient always points in the direction of steepest increase of a function, means that for an image, it always points from dark to light, and also in an image, the direction of steepest increase of intensity is typically perpendicular to an edge. I also drew one of the gradient vectors on the coordinate axes to illustrate the horizontal and vertical components of the vector. As shown on the left, those components are defined as the change of intensity in the x and y directions, the derivatives of I with respect to x and y. Let me take away the hot pink vectors to avoid clutter. We'll use the gradient to determine the direction of the motion component at a particular location like the red dot.

Next, similar to what we did in one dimension, we measure the change of intensity over time, at this location here, as the pattern moves. Depending on which direction the edge moves, the intensity at the location of the red dot will either increase or decrease over time. Finally, the quantity that we actually want to measure at this point, is the velocity of the edge in the direction along the gradient. I'll use a lower case v with the perpendicular symbol to represent this quantity, and this is a single number. If the edge is moving in the direction of the gradient, toward the light side of the edge, then this quantity will be positive. If the edge is moving in the direction opposite the gradient, toward the dark side like Darth Vader, then this quantity will be negative.

[25:51] All the quantities in items (1) and (2) here are things that we can measure directly from the image - the change of intensity in the horizontal and vertical directions in the image, and the change of intensity over time when the pattern starts moving. But this last quantity, v-perpendicular, is something we'll compute from the first two things. We'll compute it in a way that's very similar to the way we computed the x velocity in one dimension. Here we take the ratio between the time derivative and the magnitude of the gradient - the magnitude of the gradient just captures the slope of the intensity change across the edge. How do you compute the magnitude of the gradient? It's a vector, so you compute the length of the vector, by summing the squares of the x and y components, and taking the square root. If we write it out, the calculation looks like this in the denominator.

[26:58] [slide 9] So we accomplished the first goal I set for the last part of the video, measuring the perpendicular motion, which is shown here in blue. Along the way, we also determined the direction of the gradient at this location (shown with the red arrow), so we're talking about the component of motion along the direction of the red arrow. Finally, how do we express the family of velocities that are consistent with this component? We'll let (Vx, Vy) refer to the velocity of the object that we're trying to compute in the end, and the possibilities are all the vectors that have a component in the direction perpendicular here that's what we actually measured. Every one of the green vectors that I drew here, for example, has the same component in the direction perpendicular to the edge. How do we express these velocity vectors mathematically? To simplify things, we'll first construct a unit vector in the direction of the gradient - a vector in the direction of the red arrow here with a length of 1. As a quick sidebar, how would we construct a unit vector in the direction of the gradient? In the last slide, we said the gradient is defined as a vector whose two components are the derivatives of I in the x and y directions. How can we convert this to a unit vector? We divide each component by the length of the vector. The length of this vector is just the magnitude of the gradient that we wrote out at the bottom, the denominator of this expression.

[28:50] So back to our next step. Imagine that the red vector that I drew actually has a length of 1, and let (ux,uy) denote the (x,y) components of that unit vector. Then we can use the dot product between two vectors for describing the set of velocities that could have this component. A geometric interpretation of the dot product between two vectors is that it gives us the component of one vector in the direction of the other. Here we're saying that the component of the velocity vector (Vx,Vy) in the direction of the gradient (ux,uy) is the perpendicular motion that we measured, v-perpendicular. Writing out the dot product of these two vectors gives us the equation at the bottom.

[29:56] So with this in hand, we'll return to this simple picture where we were combining two measurements of the perpendicular motions to resolve the true motion of the object, assuming that it's just translating across the image, so that the velocity is the same at both locations. Now we know that from simple information we can measure directly in the image, for each location, we can compute the unit vector in the direction of the gradient and the perpendicular motion, so the things that I now wrote with green font and red font at the bottom here, these are things that we can compute from the image. The Vx and Vy shown in black are the information we want to compute in the end. We now have two linear equations in two unknowns and can solve for Vx and Vy. But, here's where I'm going to leave you hanging, we're not actually going to do this in practice, just combine two measurements of the motion components and plug them into these equations and solve for Vx and Vy. We're going to do something that leads to a solution that's more robust, so stay tuned…