Vision and Robotics

Perception Of Non-Rigid Motion: Inference of Shape, Material and Force *

Alex Pentland and John Williams

Vision Sciences Group, E15-410
The Media Lab, M.I.T.
20 Ames St., Cambridge MA 02138

Abstract

Most non-rigid motion is much simpler than is generally thought, having many fewer degrees of freedom than, for instance, 3-D shape. The relatively circumscribed nature of non-rigid motion can be seen by examining it from the perspective of *vibration modes.* The fact that a small number of parameters are sufficient to describe non-rigid motion mean that it is plausible to use sensory data to recover a nearly complete physical model of an object, so that we can, for instance, predict its response to impinging forces. One perceptual task of considerable interest is estimating material properties such as stiffness, strength, and mass by watching, touching or listening to collisions. We derive equations which allow estimates of these material properties to be extracted from visual and tactile data, and speculate on the use of aural data.

1 Introduction

Non-rigid motion seems to be very complex. When an object is struck so that it moves in a non-rigid, elastic manner, each point on the surface goes off in a different direction. It seems, therefore, that a full description of non-rigid motion must have even more degrees of freedom than a full description of shape. Despite this apparent complexity, however, people are often able to accurately predict the time course of objects experiencing non-rigid behavior, and are even able to infer material properties from observing such motion.

For instance, people can watch a tree limb swaying in the wind and make a good estimate of its stiffness, strength and mass. More commonly, when someone wants to determine the properties of an object they poke at it to see and feel it deform, and tap it to hear it resonate. From these visual, tactile, and auditory cues we are somehow able to ascertain material properties.

In this paper we will examine the finite element method of simulating the dynamics of collisions, and show that it can be decomposed into simple, closed-form differential equations describing the object's *vibration modes.* We will then show that for most objects under most conditions a relatively small number of modes are required to accurately describe how the object deforms as a function of time. Finally, we will show that for simple collisions we can estimate the object's material properties by observing these vibration modes using visual, tactile or even auditory data.

*This research was made possible by National Science Foundation Grant No. IRI-87-19920 and by ARO Grant No. DAAL03-87-K-0005

2 Background: The Finite Element Method

The finite element method (FEM) is a technique for simulating the dynamic behavior of an object. In the FEM the continuous variation of displacements throughout an object is replaced by a finite number of displacements at so-called nodal points. Displacements between nodal points are interpolated using a smooth function. Energy equations (or functionals) can then be derived in terms of the nodal unknowns and the resulting set of simultaneous equations can be iterated to solve for displacements as a function of impinging forces. In the dynamical case these equations may be written:

$$Mu + Du + Ku = f \qquad (1)$$

where u is a 3n x 1 vector of the (x, *y, z)* displacements of the n nodal points relative to the objects' center of mass, *M*, *D* and *K* are 3n by 3n matrices describing the mass, damping, and material stiffness between each point within the body, and f is a 3n x 1 vector describing the (x,y,z) components of the forces acting on the nodes. This equation can be interpreted as assigning a certain mass to each nodal point and a certain material stiffness between nodal points, with damping being accounted for by dashpots attached between the nodal points. The damping matrix *D* is normally taken to be equal to $s\backslash M + s_2K$ for some scalars s_1, s_2. When $S_1 \neq 0$ $S_2 \neq 0$ this is called Raleigh damping, for $s_2 = 0$ it is called mass damping, and for $s_1 = 0$ stiffness damping.

To calculate the result of applying some force f to the object one discretizes the equations in time, picking an appropriately small time step, solves this equation for the new u, and iterates until the system stabilizes. Direct (implicit) solution of the dynamic equations requires inversion the *K* matrix, and is thus computationally expensive. Consequently explicit Euler methods (which are less stable, but require no matrix inversion) are quite often applied.

Even the explicit Euler methods are quite expensive, because the matrices M, *D,* and *K* are quite large: for

instance, the *simplest* 3-D parabolic element produces 60 x 60 matrices, corresponding to the 60 unknowns in the 20 nodal points (x_i, y_i, z_i) which specify the element shape. In most situations M, D, and K are very much larger than 60 x 60, so that typically hundreds or thousands of very large matrix multiplications are required for each second of simulated time. For more details see references [1,2].

3 Modal Analysis

Because M, D and K are normally positive definite symmetric, and M and D are assumed to be related by a scalar transformation, Equation 1 can be transformed into $3n$ independent differential equations by use of the *whitening transform*, which simultaneously diagonalizes M, D, and K. The whitening transform is the solution to the following eigenvalue problem:

$$\lambda\phi = M^{-1}K\phi \qquad (2)$$

where λ and ϕ are the eigenvalues and eigenvectors of $M^{-1}K$.

Using the transformation $u = \phi\bar{u}$ we can re-write Equation 1 as follows:

$$\phi^T M\phi\ddot{\bar{u}} + \phi^T D\phi\dot{\bar{u}} + \phi^T K\phi\bar{u} = \phi^T f \quad . \qquad (3)$$

In this equation $\phi^T M\phi$, $\phi^T D\phi$, and $\phi^T K\phi$ are diagonal matrices, so that if we let $\bar{M} = \phi^T M\phi$, $\bar{D} = \phi^T D\phi$, $\bar{K} = \phi^T K\phi$, and $\bar{f} = \phi^T f$ then we can write Equation 3 as $3n$ independent equations:

$$\bar{M}_i\ddot{\bar{u}}_i + \bar{D}_i\dot{\bar{u}}_i + \bar{K}_i u_i = \bar{f}_i \quad , \qquad (4)$$

where $\bar{M}_i$ is the i^{th} diagonal element of $\bar{M}$, and so forth. Because the modal representation diagonalizes these matrices it may be viewed as *preconditioning* the mass and stiffness matrices, with the attendant advantages of better convergence and numerical accuracy.

What Equation 4 describes is the time course of one of the object's *vibration modes*, hence the name *modal analysis* [3]. The constant $\bar{M}_i$ is the generalized mass of mode i, that is, it describes the inertia of this vibration mode. Similarly, $\bar{D}_i$, and $\bar{K}_i$ describe the damping and spring stiffness associated with mode i, and $\bar{f}_i$ is the amount of force coupled with this vibration mode. The i^{th} row of ϕ describes the *deformation* the object experiences as a consequence of the force $\bar{f}_i$, and the eigenvalue λ_i is proportional to the natural resonance frequency of that vibration mode.

Figure 1 illustrates the some of the first and second order modes of a cylinder. Figure 1(a) shows the cylinder at rest, (b) shows the cylinder experiencing a linear deformation in response to a compressive force, (c) shows the cylinder experiencing a linear shear deformation in response to an accelerating force, (d) shows a quadratic deformation in response to a centrally-applied (bending) force, and (e) and (f) show how both the linear and second order deformations can be superimposed to produce a more accurate simulation of the object's response to the compressive and accelerating forces shown in (b) and (c).

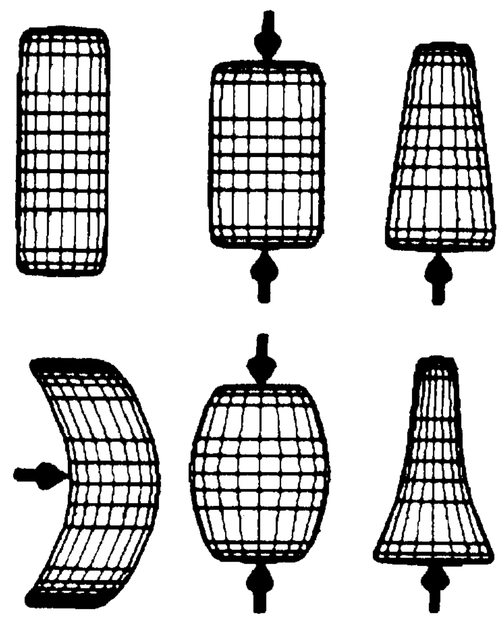

Figure 1: (a) A cylinder, (b) a linear deformation mode in response to compression, (c) a linear deformation mode in response to acceleration, (d) a quadratic mode in response to a bending force, (e) superposition of both linear and quadratic modes in response to compression, (f) superposition of both linear and quadratic modes in response to acceleration.

To obtain an accurate simulation of the dynamics of an object one simply uses linear superposition of these modes to determine how the object responds to a given force. Because Equation 4 can be solved in closed form, we have the result that for objects composed of linearly-deforming materials *the non-rigid behavior of the object in response to* an *impulse force can be solved in closed form for any time t.* The solution is discussed in Section 4. In environments with more complex forces, however, analytic solution becomes cumbersome and so numerical solution is preferred. Either explicit or implicit solution techniques may be used to calculate how each mode varies with time.

Non-linear materials may be modeled by summing the modes at the end of each time step to form the material *stress state* which can then be used to drive nonlinear plastic or viscous material behavior.

4 Number Of Modes Required

The modal representation decouples the degrees of freedom within the non-rigid dynamical system of Equation 1, but it does not reduced the total number of degrees of freedom. However once decoupled, we can separately analyze the various modes in order to determine which ones are required in order to obtain an accurate description of an object's non-rigid dynamic behavior.

The most important observation is that modes associated with high resonance frequencies normally have little effect on object shape. This is because:

- The displacement amplitude for each mode is *inversely* proportional to the square of the mode's resonance frequency. Thus higher frequencies typically have small amplitudes.

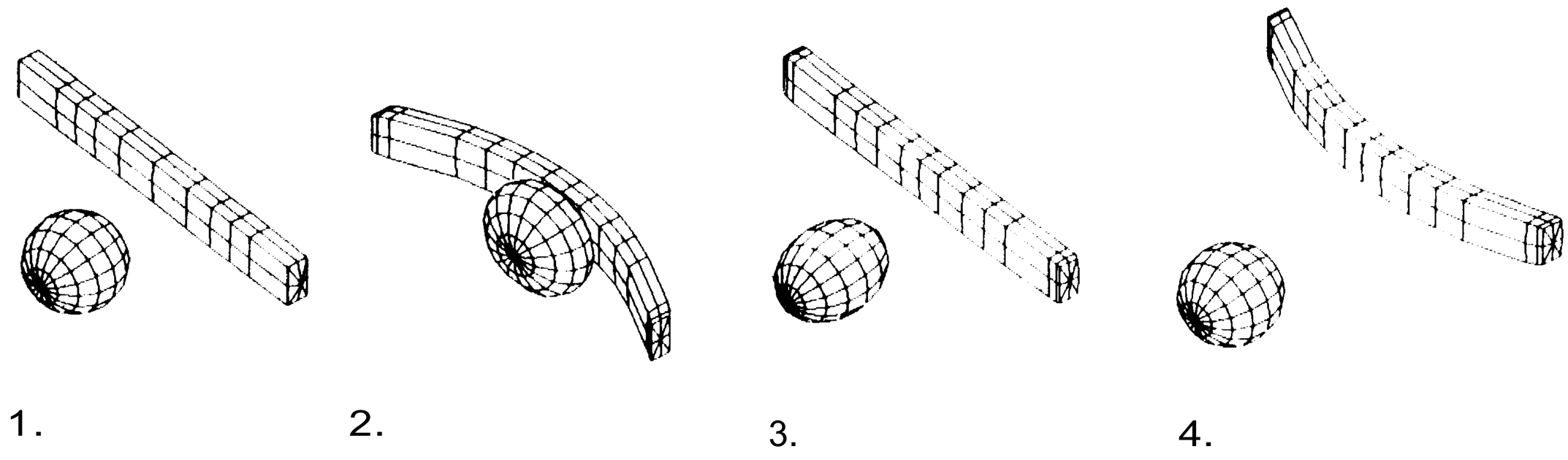

Figure 2: A ball colliding with a two-by-four

- Damping is proportional to the mode's resonance frequency. Thus higher frequency vibrations dissipate quickly.

- Frequencies are excited in proportion to the frequency content of the input force. The force generated by simple collisions typically have a roughly Gaussian time course, so that an individual resonance frequency / receives energy proportional to $_{e}$-*J 1°* Thus low frequencies typically receive more energy than high frequencies.

The combination of these effects is that high-frequency modes have very little amplitude, and even less effect, in simple collisions. As a consequence much more efficient (and still accurate) simulation of an object's dynamics can be accomplished by discarding the small, high-frequency modes, and considering only the large-amplitude, low-frequency modes [5,6]. Exactly which modes to discard can determined by examining their associated eigenvalue, which determines the resonance frequency.

Experimentally, we have found that most commonplace multi-body interactions can be adequately modeled by use of only rigid-body, linear, and quadratic strain modes. Figure 2, for instance, shows a example of a simulated non-rigid dynamic interaction: a ball colliding with a two-by-four. As can be seen, the interaction and resulting deformations look realistic despite the use of only linear and quadratic modes.[1] Figure 1 also illustrates how use of linear and quadratic modes can accurately simulate non-rigid motion. Note, however, higher-order modes are required to accurately model the objects whose dimensions differ more than an order of magnitude.

[1] Perhaps the most impressive fact about this example, however, is the speed of computation: Using a SymboUcs 3600 (with a speed of roughly one MIP), it requires only one CPU second to compute each second of simulated time!

4.0.1 Use of Fixed Modes

Normally, in either the finite element or modal methods, the mass, damping, and stiffness matrices are not recomputed at each time step. The use of fixed *M , D,* and *K* (or, equivalently, fixed modes) is well-justified as long as the material displacements are small. The definition of "small/' however, is quite different for different modes. Because the eigenvalue decomposition in Equation 2 performs a sort of principal-components analysis, it is the gross object shape (e.g., its low-order moments of inertia) determine the low-frequency modes, which as a consequence are quite stable. High-frequency modes are much less stable because they are determined by the fine features of the object's shape.

In the standard finite element formulation the action of each mode is distributed across the entire set of equations, so that one must recompute the mass and stiffness matrices as often as required by the very highest-frequency vibration modes. When these high-frequency modes are discarded the mass, damping, and stiffness matrices need to be recomputed much less frequently. In most situations it is sufficient to use single, fixed set of low-frequency modes throughout an entire simulation.

5 Estimation of Material Properties

Let us imagine that we have seen the image sequence shown in Figure 2, and have been able to compute a full 3-D description of the object's shape at each instant. How can we estimate the material properties of these objects? The first step is to compute for each object a description that decomposes the object motion into its low-order modes.

This can be accomplished by forming the vector p that describes each point on the object's surface, and then computing the matrices M, D, and *K* matrices using unit values for density and spring constant. Because these matrices are determined by the vector p up to an overall constant, the eigenvectors found by solving

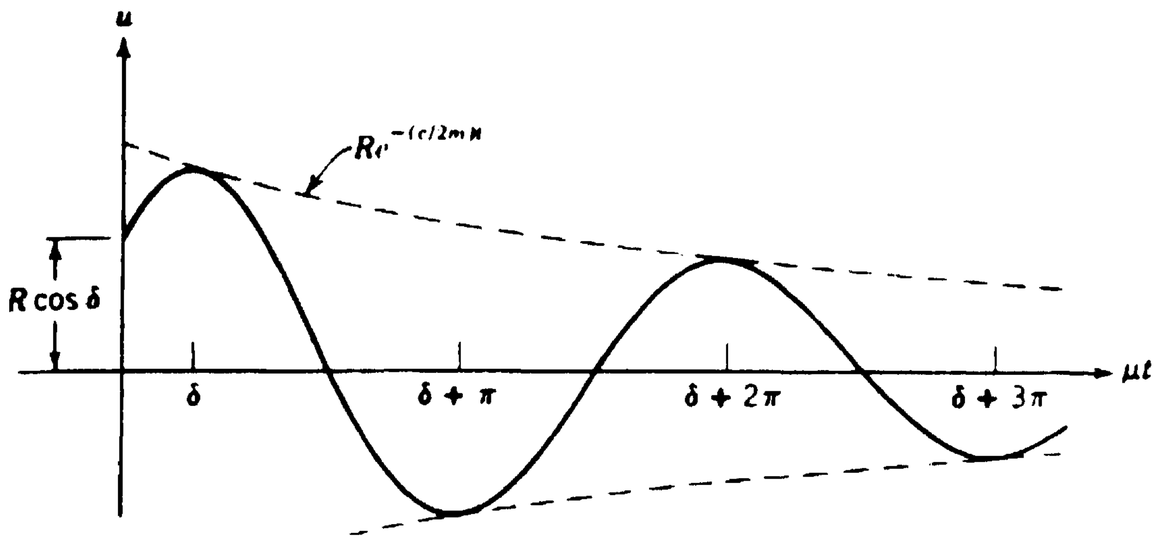

Figure 3: Plot of vibration mode amplitude versus time.

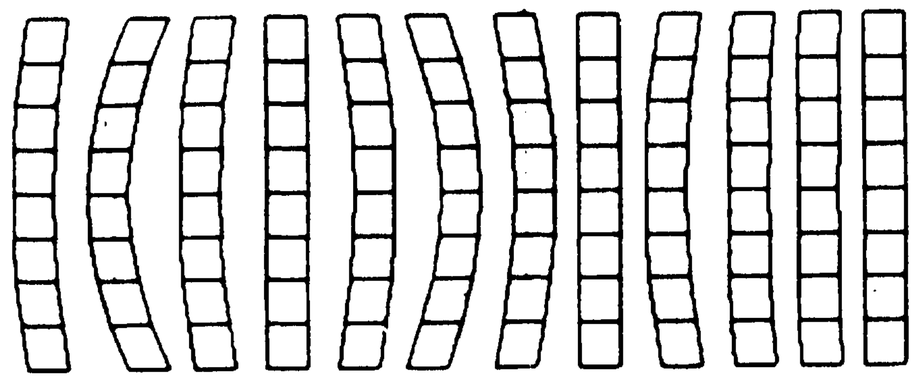

Figure 4: The two-by-four of Figure 2 vibrating after the collision, seen in canonical position.

Equation 2 will be identical regardless of the material properties (excepting, of course, non-linearly-behaving materials); only the eigenvalues depend upon the material properties. Thus given the object's shape, we can directly determine Φ, the object's deformation modes.

In this straightforward approach the matrices *M*, *D*, and *K* can become quite large, so that solution to Equation 2 becomes difficult. A more efficient method of determining Φ is be to fit either an eight or twenty point polygonal model to the sensed data (perhaps by use of the techniques described in reference [2]), and use the values of those points to construct a much smaller vector p. Given eight points we can form *M*, *D*, and *K* matrices that are 24 x 24 in size, and then solve Equation 2 to obtain the object's linear vibration modes. Given twenty points the matrices are 60 x 60, and produce both linear and quadratic vibration modes.

A still more efficient scheme can be obtained by noting that the axes of the object can allow us to directly disentangle the action of the various modes, so that by estimating axis shape and length we can avoid solving Equation 2. This approach is described in more detail below.

Given Φ and a description *p(t)* of the object's shape over time, we can directly decompose the object shape into a vector p(to) describing the object's shape at time $t = t_0$ and the amplitudes $u(t) = \Phi^{I}(p(t) - p(t_0))$ of each of the individual modes over time. The generic time behavior of each mode $u_i,(t)$ in response to an impinging force is given by Equation 4. The solution to this equation is

$$\begin{aligned}
\bar{u}_i &= Ae^{r_1 t} + Be^{r_2 t}, \\
&\qquad \text{for} \quad \bar{D}_i^2 - 4\bar{K}_i\bar{M}_i > 0, \quad r_1, r_2 < 0 \\
\bar{u}_i &= (A + Bt)e^{(\bar{D}_i/2\bar{M}_i)t}, \\
&\qquad \text{for} \quad \bar{D}_i^2 - 4\bar{K}_i\bar{M}_i = 0, \\
\bar{u}_i &= e^{(\bar{D}_i/2\bar{M}_i)t}(A\cos\mu t + B\sin\mu t), \\
&\qquad \text{for} \quad \mu = (4\bar{K}_i\bar{M}_i - \bar{D}_i^2)^{1/2}/2\bar{M}_i > 0
\end{aligned} \tag{5}$$

for the overdamped, critically damped, and underdamped cases, where

$$r_1, r_2 = \frac{-\bar{D}_i \pm \sqrt{\bar{D}_i^2 - 4\bar{K}_i\bar{M}_i}}{2\bar{M}_i} \tag{6}$$

and A and B depend on the initial conditions [11]. The third case, underdamped motion, occurs most commonly in mechanical systems and is referred to as "damped vibration." To see this we let $A = R\cos\delta$ and $B = R\sin\delta$ in Equation 5 to obtain

$$\bar{u}_i = Re^{-(\bar{K}_i/2\bar{M}_i)t}\cos(\mu t - \delta) \quad , \tag{7}$$

which is graphed in Figure 3(b).

Because the time behavior of u, is a function of M_i,-, D_i and K_i, in simple situations we can recover the physical parameters of the object's material by measuring u_i over time. This is particularly straightforward when the mode is underdamped, as in Figure 3, so that several cycles of deformation occur, because in this case we need only measure the frequency and amplitude of the deformation peaks. When the mode is critically damped or overdamped only one peak occurs; however it is still possible to measure the material properties from peaks in the first derivative of the modal amplitude.

For example, in the sequence of Figure 2 we can observe the bending of the two-by-four in response to the impact of the ball, as illustrated by Figure 4. If we plot the amplitude of this deformation mode in response to

the unknown impact force $\bar{f}_i$ (which, without loss of generality, we will take to be of unit time duration), then we will obtain the graph shown in Figure 3. From the succession of peak amplitudes $\alpha_j = \bar{u}_i(\delta + j\pi)$ and the times $t_j = \delta + j\pi$ at which these peaks occur we can obtain the following relations:

$$\bar{f}_i/\bar{M}_i = \alpha_0\pi$$

$$\bar{K}_i/\bar{M}_i = 4\pi^2(t_{j+1} - t_j)^{-2} - \tfrac{1}{4}\left(\frac{\bar{D}_i}{\bar{M}_i}\right)^2 \approx 4\pi^2(t_{j+1} - t_j)^{-2}$$

$$\bar{D}_i/\bar{M}_i = 2(\log\alpha_j - \log\alpha_{j+1})(t_{j+1} - t_j)^{-1} \qquad (8)$$

These relations are well-conditioned, so that the accuracy of the material property estimates will depend primarily on the accuracy with which we can estimate the α_j and the t_j.

Thus by observing modal amplitude after a sharp impacts we can estimate the force of collision and all of the material's physical parameters, up to a scale factor which is the mass associated with that mode. We note that except for non-linear, visco-plastic materials the relationships in Equations 8 hold exactly for *all* materials, *all* vibration modes, and *all* object geometries; they are approximations only in the sense that finite element analysis is an approximation.

5.1 Estimation of Mass

By tracking geometry over time we can estimate the material properties up to an overall scale constant that depends upon the amount of mass involved in the particular vibration mode. The obvious question, then, is how much information is required to estimate the material density? To answer this question we surveyed the CRC Handbook, which lists the density of several hundred materials. We found that the listed densities could be clustered into four categories: metals, stones, biological materials, and woody (cellulose) materials. The range of densities for metals was roughly 7.0 to 10.0, for stone materials roughly 2.3 - 2.8, for biological materials (gums, resins, soap, etc.) roughly 0.9 - 1.2, and for woody materials (e.g., common woods, plant stems, etc.) roughly 0.4 - 0.6.

Thus if the viewed material can be classified into one of these four categories, then it appears that the density can be estimated with an accuracy of roughly ±10 percent. Consequently it appears that the force, mass, damping, and spring constants can also be estimated with similar accuracy.

5.2 Static Loading

When a static load is applied to an object, a fixed deformation occurs. By examining Equation 4 we see that the amount of static deformation u_i, is a function only of the loading force f_i, and the stiffness constant K_i associated with the particular mode. Thus if we know the modes (either from analysis of the static shape or by the methods described in the next section) and the impinging force then we can determine the material stiffness K_i which is perhaps the most important of the basic material properties. In many situations, particularly when using a tactile sensor, this method of estimating material properties is probably the most practical.

5.3 Speculation: Estimation from Other Modalities

The same ideas can be applied to the tactile and auditory senses: the modal vibration can be sensed either by touch or by sound (via coupling between the air and the object surface). The notion of using a tactile sensor to produce a static measure of material stiffness has already been mentioned above. However a much stronger result can be achieved by using either tactile or auditory data given that we can apply a force f_i so as to excite single mode *i.* In this case, because we know the force, we can measure the modal mass M, directly (from Equation 8), and thus obtain D, and K_i, without knowledge of object geometry, modes, or density.

For instance, if we "poke" or "tap" an object - exert a sharp blow normal to the surface - we can transfer most of the energy into a single, localized compression mode. To the extent that we are able to isolate a single mode, we need no knowledge of object geometry, modes, or density in order to estimate the physical parameters. In experiments where the sound spectrum produced by tapping has been measured it appears that most of the energy is indeed transferred into a single mode *if* the object's geometry is uniform in the area of the tap — e.g., the object is not hollow, and the tap occurs far from any discontinuity.

6 Direct Observation of Modes

We have described a procedure of fitting data points, computing M, *D, K.* and then solving for Φ in order to obtain the object's vibration modes. Although mathematically sound, this procedure seems too complex and expensive to be useful in some applications. Similarly, the procedure is complex to be an account of human capabilities. Thus it is desirable to develop a theory that is simpler even if less accurate.

6.1 Estimation via symmetry axes

The first approach to direct observation of the modes depends on the observation that the object's axes *of symmetry* have a special status in modal analysis: they are the singular lines (and planes) of the various vibration modes, i.e., the places along which various vibration modes cause no deformation. Thus along the axes of symmetry the actions of the different vibration modes become partially decoupled.

Thus one direct way to estimate material properties is track the shape of symmetry axes over time. Time variation in the length of an axis is a function primarily of linear compression modes, while the time variation in axis curvature is primarily a function of quadratic bending modes. Either the axis length or curvature can

be tracked over time to produce a graph such as shown in Figure 3, and the material properties calculated by use of Equations 8. The simplicity of this method is its greatest virtue, however, in complex collisions the presence of many active vibration modes may destroy the accuracy of the approach.

6.2 Direct estimation of modal amplitudes

A second approach is to fit the available shape data with a simple volumetric model which can be parametrically deformed in any of the expected deformation modes. Fitting the parameters of such a model to the sensor data provides a direct, simultaneous estimate of all of the modal amplitudes. Several variations on this approach to extracting object geometry have been successfully demonstrated on range data [7,8,9,10]. As with symmetry axes, the values of such a model's deformation parameters can be tracked over time and material properties estimated by use of Equations 8. The advantage of this somewhat more computationally expensive approach is that is more likely to produce good estimates of modal amplitudes in relatively complex situations.

6.3 Estimation from static imagery

In recent years both Pentland [11] and Ley ton [12] have suggested that an object's static shape can be analyzed to produce an abstract *causal* account of the object's history. That is, that a object's shape can be decomposed into a simple prototypical shape (a cube, a sphere) that has been modified by a sequence of forces or forming operations to achieve its present state. The advantage of this type of description is that produces a prototype-based representation that can simplify object recognition [9] and which fits human psychological data [13].

It can be seen, in retrospect, that both of these theories were attempting to capture the notion that objects have a limited repertoire of modal deformations. That is, if one starts with a fixed prototypical shape (a cylinder, sphere, or cube for instance) an applies a series of forces to elastically deform that object, then in general the *simplest* possible description of the resulting shape is exactly the original shape plus the modal amplitudes that were produced by the impinging forces. Further, this description of object shape has considerable predictive power, for it contains most of the information needed to estimate material properties, determine what forces were applied, and to predict the object's future dynamic behavior.

7 Summary

Non-rigid behavior is normally simpler than it appears from cursory examination of the FEM equations that describe such behavior. By breaking such motion into vibration modes, we can show that most of the degrees of freedom have very little energy or amplitude. Thus quite accurate descriptions of non-rigid behavior can be obtained by knowing the amplitude and phase of relatively few vibration modes.

Using standard machine vision techniques the parameters of these modes can be estimated from sensory data. By tracking modal amplitudes across time (or, perhaps, even from analysis of static imagery!) we can then measure material properties, predict future behavior, and perform other ecologically important tasks.

REFERENCES

[1] Williams, J., Musto, G., and Hawking, G., (1987) The Theoretical Basis of the Discrete Element Method, *Numerical Methods in Engineering, Theory, and Application,* Rotterdam: Balkema Tubers

[2] Terzopoulos, D., Piatt, J., Barr, A., and Fleischer, K., (1987) Elastically deformable models, Proceedings of S1GGRAPII '87, *Computer Graphics,* Vol. 21, No. 3, pp 205-214.

[3] Williams, J. and Musto, G. (1987) Modal Methods for the Analysis of Discrete Systems, *Computers and Geotechnics,* Vol. 4, pp 1-19.

Pentland, A., and Williams, J. (1988) Virtual Construction, *Construction,* Vol. 3, No. 3, pp. 12-22

[5] Pentland, A. and Williams, J., (1989) Good Vibrations: Modal Analysis for Graphics and Animation, Proceedings of SIGGRAPH '89, *Computer Graphics,* Vol. 23, No. 3.

[6] Pentland, A., and Williams, J. (1989) Fast Simulation on Small Computers: Modal Dynamics Applied to Volumetric Models *IEEE 20th Annual Conference on Simulation,* May 4-5, Pittsburgh, PA.

[7] Pentland, A., (1988) Part segmentation for object recognition, *Neural Computation,* Vol. 1, No. 1

[8] Pentland, A., (1988) Automatic extraction of deformable part models, *M.I.T. Media Lab Vision Sciences Technical Note 104,* July 1, 1988.

[9] Bajcsy, R. and Solina, F., (1987) Three-dimensional Object Representation Revisited, *First International Conf. on Computer Vision, '87,* June 8-11, London, England.

[10] Bolt, T., and Gross, A., (1987) Recovery of Superquadrics from Depth Information, *Proceedings of the AAA1 Workshop on Spatial Reasoning and Multi-Sensor Fusion,* Oct. 1987, pp. 128-137

[11] Pentland, A. (1986) Perceptual Organization and the Representation of Natural Form, *Artificial Intelligence Journal,* Vol. 28, No. 2, pp. 1-38.

[12] Ley ton, M. (1988) A Process-Grammar for Shape, *Artificial Intelligence Journal,* No. 34, pp. 213-247.

[13] Rosch, E. (1973) On the internal structure of perceptual and semantic categories. In *Cognitive Development and the Acquisition of Language.* Moore, T.E. (Ed.) New York: Academic Press.