WORKSHOP SOLUTION - POINT 8

What strategies do you know (or can think of) in order to make PCA more robust?

Principal Component Analysis (PCA) is the problem of finding a low rank approximation to a matrix. It is a central problem in statistics, but it is sensitive to sparse errors with large magnitudes. Robust PCA addresses this problem by decomposing a matrix into the sum of a low-rank matrix and a sparse matrix, thereby separating out the sparse errors.

Traditional PCA can handle small noise but is brittle with respect to grossly corrupted observations-- even one grossly corrupt observation can significantly mess up answer.

Robust PCA factors a matrix into the sum of two matrices, M=L+S, where M is the original matrix, L is low-rank, and S is sparse. This is what we'll be using for the background removal problem.

Low-rank means that the matrix has a lot of redundant information-- in this case, it's the background, which is the same in every scene (talk about redundant info!). Sparse means that the matrix has mostly zero entries-- in this case, see how the picture of the foreground (the people) is mostly empty. (In the case of corrupted data, S is capturing the corrupted entries).