Chapter 4

Taxonomy of Asteroids

This chapter describes the Bus-DeMeo taxonomy which began as a Masters thesis (DeMeo, 2007), but was expanded and solidified into published form as part of this PhD thesis. Much of the text in Sections 4.1 to 4.3 has been taken directly from DeMeo et al. (2009a). In this chapter I focus on describing the taxonomy and analyzing what has been learned about the usefulness of visible and near-infrared data in spectral analysis. For more discussion of the principal component boundaries for the classes, the reader is referred to DeMeo (2007) and DeMeo et al. (2009a).

Contents

4.1 Need for a new taxonomy

Taxonomic classification systems for asteroids have existed since there were enough data to distinguish meaningful groups. The first taxonomies were based on asteroid broad band filter colors such as Wood and Kuiper (1963) and Chapman et al. (1971) where they noted two separate types of objects denoted as "S"and "C". Figure 4.1 shows the separation of the C and S classes using U-B and B-V colors from Bowell et al. (1978). Taxonomies and their nomenclature grew and evolved as later taxonomies became based on higher resolution spectral data which reveal features offering clues to surface composition, age, and alteration. The most widely used taxonomies for asteroids currently are the Tholen taxonomy (1984) based on the Eight-Color Asteroid Survey data (ECAS, Zellner et al., 1985) and SMASSII spectral taxonomy (Bus, 1999; Bus and Binzel, 2002b,a) based on the SMASSII spectral dataset. The average spectra for each of their classes are shown in Figures 4.1 and 4.2. For a review of the evolution of asteroid taxonomies see Bus (1999).

Both the Tholen and Bus taxonomies were based on Principal Component Analysis, a dimension-reducing technique first applied to the field of asteroid classification by Tholen (1984). Most previous asteroid taxonomies were based on visible data because only in the current decade has spectral data collection become widely available in the near-infrared for asteroids down to relatively faint (V=17) limiting magnitudes. The instrument SpeX on the NASA Infrared Telescope Facility (IRTF) has been crucial to increasing the library of near-IR asteroid spectra. (Rayner et al., 2003)

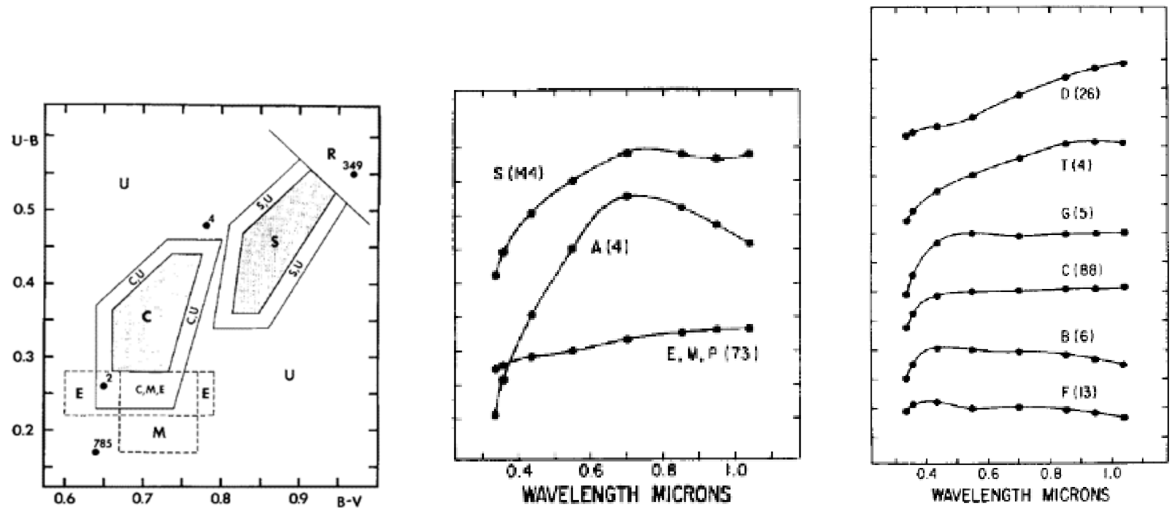

Figure 4.1: **Left:** Separation of classes in Bowell Taxonomy using U-B and B-V colors Bowell et al. (1978). **Middle and Right:** Example spectra for each class in the Tholen taxonomy (Zellner et al., 1985; Tholen, 1984)

The near-IR data range reveals diagnostic compositional information because of the presence of features at one and two microns primarily due to the presence of olivine and pyroxene. Other classification systems created using near-IR data include Howell et al. (1994) who created a neural network taxonomy. Gaffey et al. (1993) created an S-complex taxonomy of olivine- and pyroxene-rich asteroids based on near-infrared data. Our goal was to create a taxonomy extending from visible to near-infrared wavelengths for the entire suite of asteroid characteristics with a method that can be easily reproduced by future users to classify new data. We also strove to keep the notation consistent with past taxonomies, specifically the Bus taxonomy, to facilitate the transition to this new system. The Bus taxonomy, in turn, strove to keep its notation consistent with the Tholen taxonomy.

The taxonomy was created using Principal Component Analysis which is described in detail in Section 3.1.2. It is comprised of 24 classes compared to 26 in the Bus system with three Bus classes eliminated (Sl, Sk, Ld) and one (Sv) created, as well as the addition of a "w" notation, a mark meant to flag objects

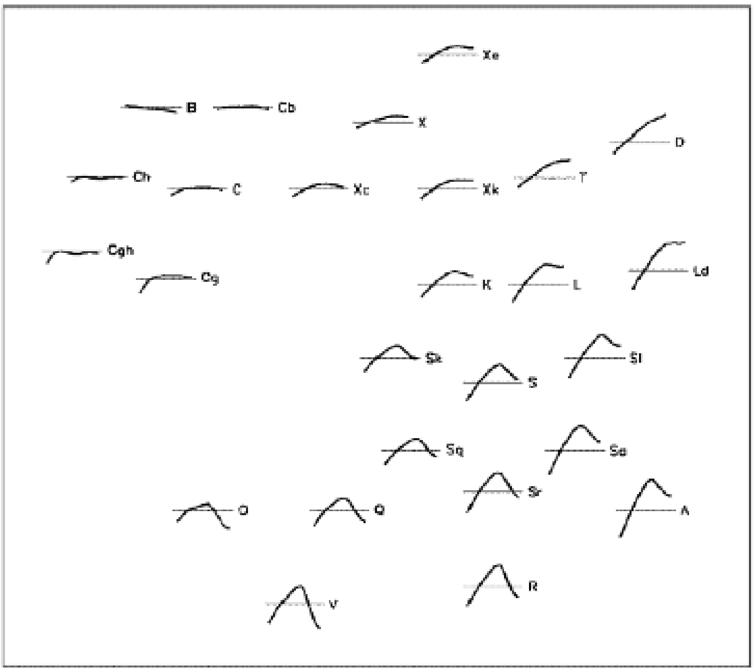

Figure 4.2: A key of the 26 spectral classes of the Bus taxonomy mapped out by their locations in principal component space (Bus and Binzel, 2002a)

having similar spectral features but differing only by having a higher spectral slope.

The taxonomy classes are formally defined by data spanning the wavelength range 0.45 to 2.45 microns as compared with 0.34 to 1.04 microns for Tholen (1984) using eight points, and 0.435 to 0.925 microns for Bus (1999) using 48 points. A method of interpreting near-infrared data from 0.85 to 2.45 microns is also described but for many classes IR-only data do not yield a unique outcome in Principal Component Analysis (PCA) and the data cannot formally be classified.

4.2 The Data

The new data used for creating this taxonomy are near-infrared spectral measurements from 0.8 to 2.5 microns obtained using SpeX in single prism mode (R=250) with a slit width of 0.8 arcseconds.

As described in DeMeo and Binzel (2008), objects and standard stars were observed near the meridian to minimize their differences in airmass and match their parallactic angle to the fixed N/S alignment of the slit. Frames were taken so that the object was alternated between two different positions (usually noted as the A and B positions) on a 0.8 x 15 arcsecond slit aligned north-south. The asteroid spectrum was divided by the spectrum of a solar-type star, giving relative reflectance. Our primary solar analog standard stars were 16 Cyg B and Hyades 64. Additional solar analog stars with comparable spectral characteristics were utilized around the sky. Two to three sets of eight spectra per set were taken for each object, with each with exposures typically being 120 seconds. The total integration time for each of these objects therefore ranged from 30 to 120 minutes.

Reduction was performed using a combination of routines within the Image Reduction and Analysis Facility (IRAF), provided by the National Optical Astronomy Observatories (NOAO) (Tody, 1993), and Interactive Data Language (IDL). We use a software tool called “Autospex” to streamline reduction procedures. Autospex writes macros containing a set of IRAF (or IDL) command files that are then executed by the image processing package. Autospex procedures operate on a single night at a time, with the opportunity for the user to inspect and verify the results at each stage. Briefly, autospex writes macros that: trim the images down to their useful area, create a bad pixel map from flat field images, flat field correct all images, perform the sky subtraction between AB image pairs, register the spectra in both the wavelength and spatial dimensions, co-add the spectral images for individual objects, extract the 2-D spectra from co-added images, and then apply the final wavelength calibration.

Using IDL, an absorption coefficient based on the atmospheric transmission (Atmospheric Transmission Model (ATRAN)) model by Lord (1992) is determined for each object and star pair that best minimizes atmospheric water absorption effects for that pair. This coefficient correction is most important near 1.4 and 2.0 microns, locations of major absorption bands due to telluric H_2O. The final IDL step averages all the object and standard pairs to create the final reflectance spectrum for each object.

Most (321) visible wavelength spectra (usually 0.4 to 0.9 microns) were taken from the Small Main-Belt Asteroid Spectroscopic Survey (SMASS) data set (Bus and Binzel, 2002b). Our sample was comprised of 371 objects with both visible and near-IR data.

4.3 The Taxonomy

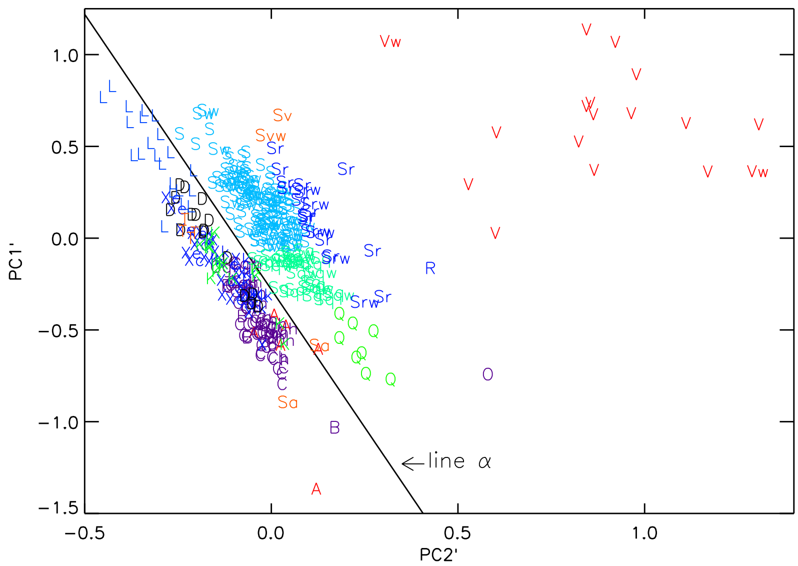

Figure 4.3: Results for PC2′ versus PC1′. All objects plotted are labeled with their taxonomic classification in this system. Notice the "grand divide" between the S-complex and the C- and X-complexes. Line α separates objects with and without 2-μm absorption bands. The direction orthogonal to line α (increasing PC2′ values) indicates deeper 2-μm and narrower 1-μm absorption bands. The direction parallel to line α (increasing PC1′ values) indicates wider 1-μm absorption bands. The notation "PC1′", "PC2′", etc. denotes that these principal components are computed after removal of the slope.

The guiding principle for the classification rules of this taxonomy was to define regions of principal component space that most consistently envelop objects within each of the original Bus (1999) classes. With this principle as a guide, we subjectively define boundaries so that the most similar spectra consistently fall into the same taxonomic classes. The over-riding criterion of similarity of spectral properties in a class, as examined over the full 0.45 - 2.45 μm range, led to some objects in the Bus (Bus, 1999; Bus and Binzel, 2002a) classification receiving new class designations.

Of the 371 objects in our sample, 321 were previously assigned labels within the Bus taxonomy. We used this set of 321 objects to guide the class boundaries. Details of the process and descriptions of the boundaries created are explained thoroughly in DeMeo et al. (2009a) and the flow charts outlining each step to classification are listed Appendces B and C of that work. In Figs. 4.3 and 4.4, the main

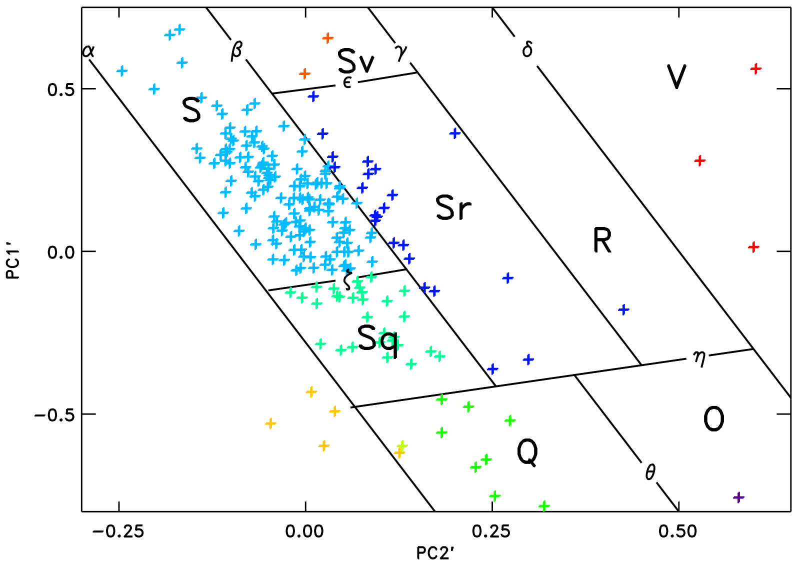

Figure 4.4: PC2′ versus PC1′ plotted for the S-complex plus A-, Q-, O-, R-, and V-types. Boundaries chosen for each class are shown and lines are labeled with greek letters. All boundaries are perpendicular or parallel to line α. A- and Sa- types are the only classes which can lie on either side of line α.

results of the principal component analysis are shown in principal component space. Particularly, Fig. 4.3 displays the "grand divide" (labeled as line α) that separates the featured and subtly featured spectra, and Fig. 4.4 shows the break up of the S-complex.

It is important to note that throughout this chapter on taxonomy, the most basic division among spectra is between "featured" and "subtly featured" spectra. By "featured" we mean the spectrum contains a prominent 1 or 2 micron band. By "subtly featured" we mean there may or may not be shallow absorption features particularly in the visible wavelength range, however, there are no prominent one or two-micron absorption bands.

The taxonomy is comprised of 24 classes (compared to the 26 in Bus and 8 in Tholen). Some argue that too many classes make taxonomy and classification confusing, however, without the proper level of detail compared to the quality of the data, a taxonomy is of little use. For those who are not spectroscopists and seek a "simpler" system, the taxonomy has a hierarchy that suits their needs. The taxonomy, consistent with previous work, has three main "complexes" that encompass the large majority of all spectral types. These include the S-complex, C-complex, and X-complex. This notation has existed nearly since the invention of asteroid taxonomies and puts classification into its most simple form. Within those complexes are classes which subdivide the spectra in further detail. In addition to the complexes, there are the "end members." Those classes represent fairly unique spectra for which there is not a very large sample. Thus this taxonomy can suit the needs of those who seek simplicity as well as those who need more of the detail contained within a spectrum. Here we present the results of the analysis by describing the characteristics of each class over the visible and near-IR wavelengths (and largely overlook the details of PCA). For a table of observations and references for all data included in this work as well as the final taxonomic designations for all objects see Appendix A.1. The spectra are plotted in Appendix A.2.

4.3.1 The end members: A, V, R, O, Q

Spectrally, the A-class has a deep and extremely broad absorption band with a minimum near 1 μm and may or may not have shallow 2-μm absorption band; it is also steeply sloped. It is spectrally very unique and therefore easily identifiable. Figure 4.5 shows the spectral progression from S to A (the S-and Sa-types are described in section 4.3.2).

The V-class, based on the asteroid 4 Vesta (Tholen 1984), is characterized by its strong and very narrow 1-μm absorption band, as well as a strong and wider 2-μm absorption feature. Most V-class asteroids that have been discovered are among the Vesta family and are known as Vestoids, although a few other objects have been identified throughout the main belt, such as 1459 Magnya (Lazzaro et al., 2000) and objects from the basaltic asteroid survey by Moskovitz et al. (2008).

The R-class, created for its sole member 349 Demboska by Tholen (1984), is similar to the V-class in that it displays deep 1- and 2-μm features, however the one-micron feature is broader than the V-type feature and has a shape more similar to an S-type except with deeper features. Bus (Bus and Binzel, 2002a) included three other members in the R-class, two of which are included in our sample. These two objects (1904 Massevitch and 5111 Jacliff) were reassigned to the V-class after discovering that in the near-infrared their one-micron bands remain very narrow. Moskovitz et al. (2008) list 5111 Jacliff as an "R-type interloper" within the Vesta family, but it appears to be an object more confidently linked to Vesta. 1904 Massevitch, however, has a semi-major axis of 2.74 AU. The unusual spectrum and outer belt location for asteroid 1904 has been noted previously (e.g. Burbine and Binzel (2002)). In the sample we present here, asteroids 1904 Massevitch and 1459 Magnya (Lazzaro et al., 2000) are the only two V-types beyond 2.5 AU, a region where V-type asteroids are rare (Binzel et al., 2006a, 2007; Moskovitz et al., 2008).

The O-class also has only one member, 3628 Boznemcova, defined by Binzel et al. (1993). Boznemcova is unique with a very rounded and deep, bowl shape absorption feature at 1 micron as well as a significant absorption feature at 2 μm. Even though the class is separated in the flow chart, more data on R-type and O-type objects may help establish more rigorously their region boundaries. Bus (Bus and Binzel, 2002a) designated three other asteroids as O-type, 4341 Poseiden, 5143 Heracles, and 1997 RT. Only 5143 was included in our sample. Asteroid 5143 is reclassified here as a Q-type because with near-infrared data it is clear the object did not have the distinct "bowl" shape of the one-micron feature of Boznemcova. This adds 5143 Heracles as a Q-type to those known within near-Earth space (e.g., Binzel et al., 2004).

The Q-class, whose boundaries are labeled in Fig. 4.4 was first defined by Tholen (1984) for near-Earth asteroid 1862 Apollo. The class is characterized by a deep and distinct 1-μm absorption feature with evidence of another feature near 1.3 μm as well as a 2-μm feature with varying depths among objects. The spectral differences between the end member classes A, V, R, Q, and O are displayed in Figure 4.5.

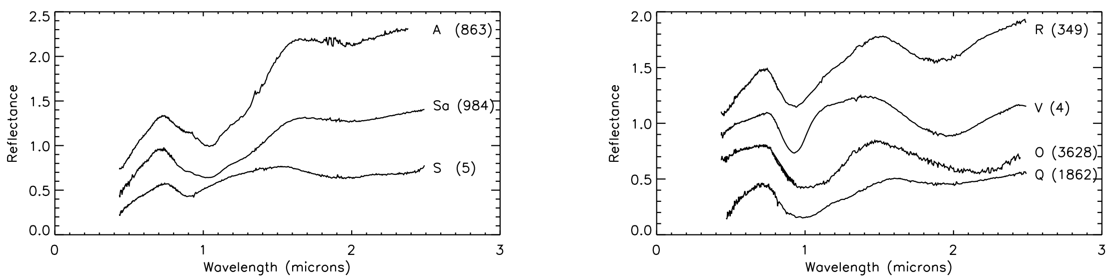

Figure 4.5: **Left:** Examples of S-, Sa-, and A-classes. There is a clear progression from S-types with a shallow one-micron band and low slope to A-types with a deep one-micron band and high slope. Sa- and A-types show similar 1-μm band absorptions, but Sa-types are much less red than A-types. The class and the asteroid number are labeled next to each spectrum. **Right:** Comparison of prototypes for the V-, O-, Q-, and R-classes. Note the O-class has a very wide 1-micron band and the V-class has a very narrow band. The V-types with the deepest 2-micron bands plot farthest from line α. For this and all subsequent spectral plots: We present relative reflectances normalized to unity at 0.55 microns; the spectra are offset vertically for clarity of comparison. The class and asteroid number are labeled next to each spectrum.

4.3.2 The S-complex: S, Sa, Sq, Sr, Sv

Just as in the case of the Bus taxonomy, the S-complex was by far the most difficult to subdivide. Most Bus classes within the S-complex seemed to blend together or scatter randomly in all combinations of PCA components. For example, many objects labeled as "Sa", "Sl", and "Sk" in the Bus (Bus, 1999; Bus and Binzel, 2002a) taxonomy no longer form distinct groups when their spectra are extended into the near-IR. Most original Bus class objects of these types merged into the S-class. The Sl and Sk classes were excluded from this new system. The Sa-class was kept, however, it was redefined and no longer contains any of the objects previous designated by the Bus system. Similarly, many Bus S, Sq, and Sk objects become less clearly separated when their spectra extend to the near-infrared. Within PCA space, the Bus S, Sq, and Sk objects were initially impossible to define clearly because the boundaries blur and overlap. Because spectrally the main difference between the classes of the S-complex appears to be the width of the 1-micron absorption band we used the wavelength range 0.8 to 1.35 microns and performed PCA on only S-complex objects to gain insight on their differences. This S-class PCA served as a guide to separate S- complex classes in a meaningful way based on the near-infrared spectral features.

The Sa-class, the most distinct among all S-complex types, has the same characteristic 1-μm absorption band as the A-class, but is less red. A figure showing the spectral progression from S to Sa to A is shown in Fig. 4.5. The current Sa-class was redefined from the Bus system because the two Sa objects (main belt object 984 Gretia and Mars crosser 5261 Eureka) in this system were both Sr-types in the Bus system. Since these objects prove to be intermediate between S and A we change the classification of these two (Bus) Sr-types to Sa in this taxonomy.

The S-class has moderate 1- and 2-μm features. The Sq-class has a 1-μm absorption band that is wider than that of an S-type with evidence of a feature near 1.3 μm like the Q-type, except the 1-μm feature is more shallow for the Sq. Many objects that were previously designated as Sa-, Sl-, or Sq- types in the Bus taxonomy were designated as S- or Sq-types in this extended taxonomy.

The Sr-class typically has a fairly narrow 1-μm feature similar to but more shallow than an R-type as well as a 2-μm feature. One object (5379 Abehiroshi) was a V-type under the Bus (Bus and Binzel, 2002a) system and is now labeled an Sr. While the visible data have a "moderate to very steep UV slope shortward of 0.7 μm with a sharp, extremely deep absorption band longward of 0.75 μm" (Bus and Binzel, 2002a), it is clear with the inclusion of near-infrared data that the one-micron absorption band is too wide to be a V-type.

The Sv-class has a very narrow 1-μm absorption band similar to but more shallow than a V-type as well as a 2-μm feature. Two objects (2965 Surikov and 4451 Grieve) are considered spectrally unique from Sr because they exhibit very narrow 1-μm absorption bands. The objects in this region spectrally appear to be in transition between S- and V-classes. They are not included in the Bus dataset, and Bus and Binzel (2002a) did not report any cases of objects with these characteristics. Because of their intermediate properties between S and V that are clearly displayed over the 0.45- to 2.45-micron range, we define a new class with the label Sv. Figure 4.3.2 displays the spectra of typical S-, Sq-, Sr-, and Sv-class spectra.

4.3.3 The w-notation

The objects in the S-complex had widely varying spectral slopes. To have some taxonomic distinction in spectral characteristics arising from slope, we made an arbitrary cutoff at Slope = 0.25 dividing high slope objects from other objects. These objects are not relabeled in a class of their own. Instead the S, Sq, Sr, and Sv objects with high slopes receive a notation of w added to their name as a moniker for what is commonly discussed as an increase in slope arising from space weathering (Clark et al., 2002). [We make no pretense of knowing whether or not their surfaces are actually weathered.] The high slope S objects are labeled Sw, Sqw, Srw and Svw. We extended this flag to the V- types for which there were two objects with slopes greater than 0.25, which we label as Vw. Sa-types do not receive a w notation because, as an intermediate class between S and A, they are by definition highly sloped. Figure 4.3.2 displays the differences between low- and high-slope objects, S and Sw.

The choice of 0.25 for the "w" notation is arbitrary. When plotting Bus labeled S, Sa, and Sl objects, there is a mixing around the 0.23 to 0.27 slope range. The goal was to keep the "w" notation more selective without setting the boundary too high where objects with unusual slope features (such as deeper UV

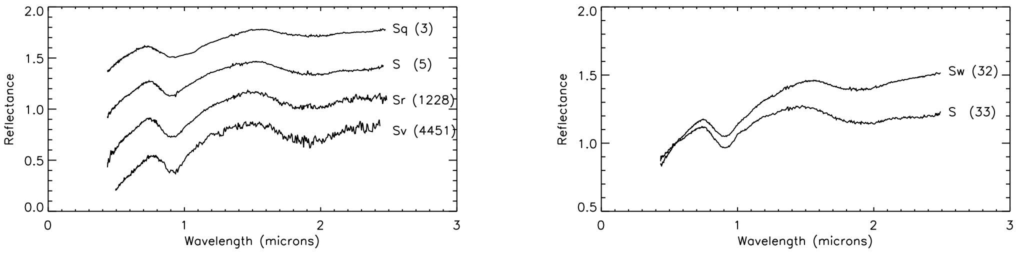

Figure 4.6: **Left:** Comparison of spectra within the S-complex (S, Sq, Sr, Sv) showing the variation in the one-micron absorption band among these types. Sq-types have the widest one-micron feature, similar to the Q-class. Sv-types have the narrowest feature, similar to the V-class. **Right:** Illustration of S and Sw reflectance spectra. The absorption features for both are very similar. Slope is the most significant distinction between the two, where the "w" is a notation to denote the slope difference, but does not describe a distinct class. These two spectra are not offset vertically, showing their differences relative to their common normalization at 0.55 μm.

dropoffs) were preferentially selected rather than focusing on the significant slope range between one and two microns for the S-Complex.

4.3.4 The end members: D, K, L, T

The D-class spectra are linear with a very steep slope, and some show slight curvature or a gentle kink around 1.5 μm. The T-class is linear with moderate to high slope and often gently concaving down, It is preserved from the Bus system, although it is very similar to the X-class in the near-infrared.

The Bus (Bus and Binzel, 2002a) L- and K-classes were part of the S-class in the Tholen (1984) taxonomy. While the L-class may show 1- and 2-μm features, it is distinct from the S-class because the steep slope in visible region levels out abruptly around 0.7 μm, but does not show a distinct absorption band like the S. There is often a gentle concave down curvature in the near-infrared with a maximum around 1.5 μm, and there may or may not be a 2-micron absorption feature. A typical K-class object displays a wide absorption band centered just longward of 1 μm. This feature is unique because the left maximum and the minimum are sharply pointed and the walls of the absorption are linear with very little curvature. Figure 4.7 shows examples of typical spectra in the D-, K-, L- and T- classes.

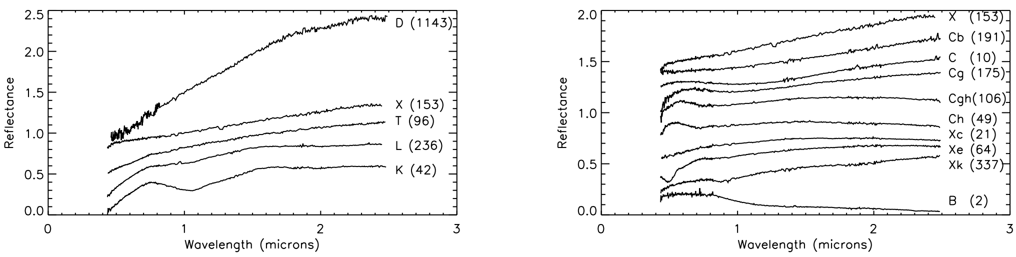

Figure 4.7: **Left:** Prototype examples of D-, L-, K-, T-, and X-class spectra. **Right:** Prototype examples for C- and X-complex spectra.

4.3.5 C- and X- Complexes: B, C, Cb, Cg, Cgh, Ch, X, Xc, Xe, Xk

Over the wavelength range used for this work, PCA is not particularly sensitive to the subtle features that define the C and X complexes. This taxonomy generally strives to follow the definitions created by Bus and Binzel (2002a) because most features exist in the visible wavelength range. Significant analysis was performed to distinguish these classes in the near-IR, which is discussed further in Section 4.6. Figure 4.3.4 shows typical spectra for classes within the C- and X-complexes.

The B-types are easily distinguished by their negative slope. Their spectra are linear and negatively sloping often with a slight round bump around 0.6 μm preceding a slight feature longward of 1 micron and/or a slightly concave up curvature in the 1- to 2-μm region.

Cb-types are linear with a small positive slope that begins around 1.1 μm. Cb objects were intermediate objects between the B- and C-classes in the Bus system (Bus and Binzel, 2002a). We keep the same notation, however, the near-infrared data shows, that Cb objects have low to moderate near-infrared slopes, while the visible slopes are low or negative.

C-types are linear with neutral visible slopes and often have a slight rough bump around 0.6 μm and low but positive slope after 1.3 μm. They also may exhibit slight feature longward of 1 μm. Ch spectra have a small positive slope that begins around 1.1 microns and slightly pronounced UV dropoff, and a broad, shallow absorption band centered near 0.7 μm. The Cgh-class is similar to the Ch showing a 0.7-micron feature, but also has a more pronounced UV dropoff like the Cg-type.

There is only one object (175 Andromache) in the Cg-class carrying over from the Bus (Bus, 1999; Bus and Binzel, 2002a) taxonomy. The Cg-class is characterized by a pronounced UV dropoff similar to the Cgh, but does not show the 0.7-micron feature that define Ch and Cgh.

The X-class is identified based on medium to high slope values and its linear spectrum. Xc-types have low to medium slope and are slightly curved and concave downward. The Xe-class exhibits low to medium slope similar to either Xc- or Xk-type, but with an absorption band feature shortward of 0.55 μm. The Xk-class is slightly curved and concave downward similar to Xc-type but with a faint to feature between 0.8 to 1 μm. The spectral slope after this feature varies widely among spectra.

A summary of the descriptions of each spectral class is provided in Table 4.1. A key of the taxonomy is plotted as the average spectrum for each class in Figs. 4.8 and 4.9.

Table 4.1: Spectral Class Descriptions

Class	**Description**	**Prototypes**
A	Deep and extremely broad absorption band with a minimum near 1 μm, may or may not have shallow 2-μm absorption band; very highly sloped.	246, 289, 863
B	Linear, negatively sloping often with a slight round bump around 0.6 μm and/or a slightly concave up curvature in the 1- to 2-μm region.	2, 3200
C	Linear, neutral visible slope often a slight rough bump around 0.6 μm and low but positive slope after 1.3. May exhibit slight feature longward of 1 μm.	1, 10, 52
Cb	Linear with a small positive slope that begins around 1.1 μm.	191, 210, 785
Cg	Small positive slope that begins around 1.3 microns and pronounced UV dropoff.	175
Cgh	Small positive slope that begins around 1 micron and pronounced UV dropoff similar to Cg also includes a broad, shallow absorption band centered near 0.7 μm similar to Ch.	106, 706, 776
Ch	Small positive slope that begins around 1.1 microns and slightly pronounced UV dropoff also includes a broad, shallow absorption band centered near 0.7 μm.	19, 48, 49
D	Linear with very steep slope, some show slight curvature or gentle kink around 1.5 μm.	1143, 1542, 3248
K	Wide absorption band centered just longward of 1 μm, the left maximum and the minimum are sharply pointed and the walls of the absorption are linear with very little curvature.	42, 579, 742
L	Steep slope in visible region leveling out abruptly around 0.7 μm. There is often a gentle concave down curvature in the infrared with a maximum around 1.5 μm. There may or may not be a 2-micron absorption feature.	236, 402, 606
O	Very rounded and deep, bowl shape absorption feature at 1 micron as well as a significant absorption feature at 2 μm.	3628
Q	Distinct 1-μm absorption feature with evidence of another feature near 1.3 μm; a 2-μm feature exists with varying depths between objects.	1862, 3753, 5660
R	Deep 1- and 2-μm features; the one-micron feature is much narrower than a Q-type, but slightly broader than a V-type.	349
S	Moderate 1- and 2-μm features. The 2-micron feature may vary in depth between objects.	5, 14, 20
Sa	Has a deep and extremely broad absorption band at 1 μm; has similar features to A-types but is less red.	984, 5261
Sq	Has a wide 1-μm absorption band with evidence of a feature near 1.3 μm like the Q-type, except the 1-μm feature is more shallow for the Sq.	3, 11, 43
Sr	Has a fairly narrow 1-μm feature similar to but more shallow than an R-type as well as a 2-μm feature.	237, 808, 1228
Sv	Has a very narrow 1-μm absorption band similar to but more shallow than a V-type as well as a 2-μm feature.	2965, 4451
T	Linear with moderate to high slope and often gently concaving down.	96, 308, 773
V	Very strong and very narrow 1-μm absorption and as well as a strong 2-μm absorption feature.	4, 1929, 2851
X	Linear with medium to high slope.	22, 87, 153
Xc	Low to medium slope and slightly curved and concave downward.	21, 97, 739
Xe	Low to medium slope similar to either Xc- or Xk-type, but with an absorption band feature shortward of 0.55 μm.	64, 77, 3103
Xk	Slightly curved and concave downward similar to Xc-type but with a faint to feature between 0.8 to 1 μm.	56, 110, 337
Bus-Ld	Diverged to L- and D-classes.	279 (D), 3734 (L)
Bus-Sk	Diverged to the S- and Sq-classes.	6585 (S), 3 (Sq)
Bus-Sl	Merged with the S-class.	17 (S), 30 (S)

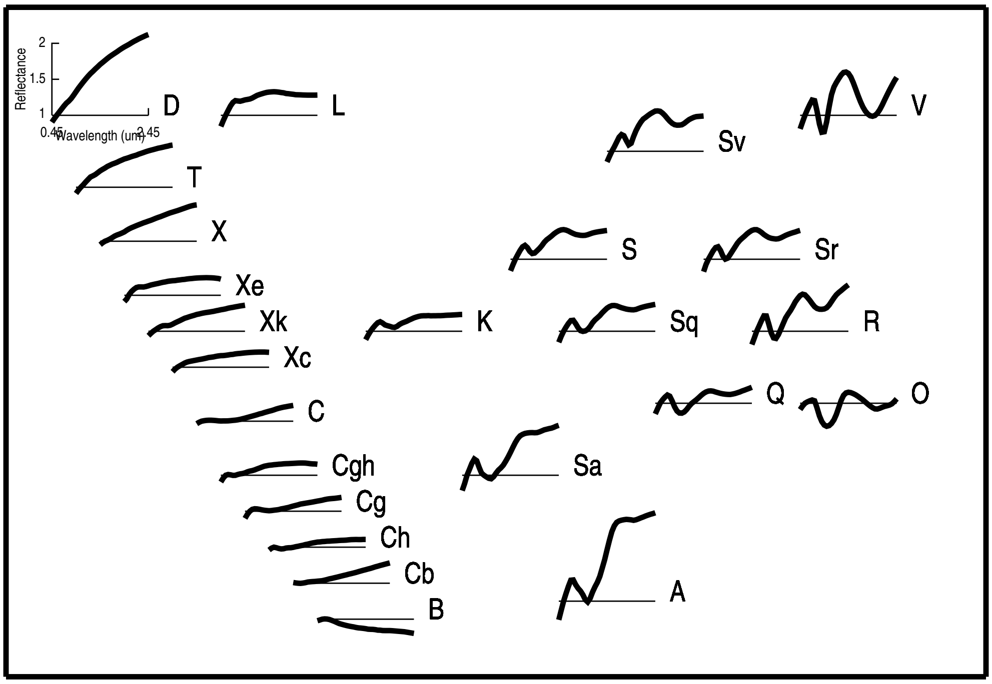

Figure 4.8: A "key" showing all 24 taxonomic classes defined over 0.45 - 2.45 μm. The average spectra are plotted with constant horizontal and vertical scaling and are arranged in a way that approximates the relative position of each class in the primary spectral component space defined by slope, PC1′, and PC2′. Thus the depth and width of the 2-μm band generally increases lower left to upper right, and the depth and width of the 1-μm band increase moving downward and toward the right. For subtly featured objets slope increases from bottom to top. Due to the spectral complexity of the C- and X-complexes, the locations of some of these classes do not strictly follow the pattern. The horizontal lines to which each spectrum is referenced has a reflectance value of unity. This figure and description follows the style of Bus and Binzel (2002a, Fig. 15).

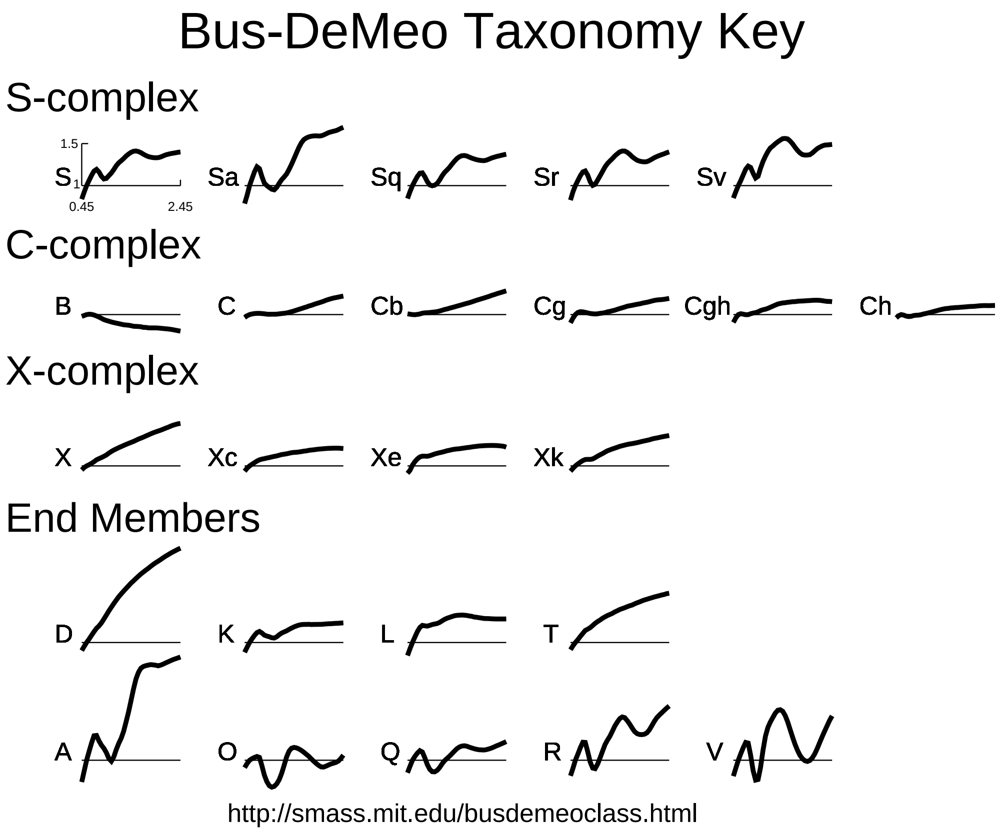

Figure 4.9: Re-arranged key to Bus-DeMeo taxonomy showing average spectra for each of the 24 classes, grouped according to their complex.

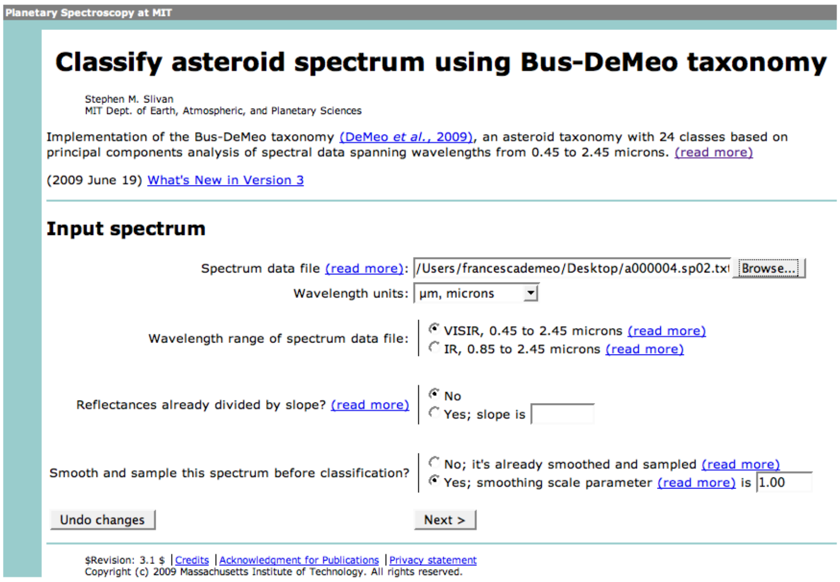

Figure 4.10: Taxonomy Web Tool Application Step 1: Input Spectrum. Here the user specifies the wavelength units, the wavelength range, and whether or not the slope has been removed and sample has been smoothed.

4.4 Taxonomy Web Application

A web application (constructed by Dr. Stephen Slivan) was created that determines taxonomic types for visible plus near-infrared data or near-infrared-only data based on this extended taxonomy (smass.mit.edu). This tool takes an input spectrum and creates a spline fit with the appropriate data values. It then calculates the principal component values, and classifies the spectrum using the appropriate (visible+near-IR or near-IR-only) flow chart. The final output displays the classification, the principal component values, a definition of the class (or classes if there are multiple possibilities), and plots the spectrum with the average spectrum from each possible class along with the residual to give an quantitative estimate of how good the fit is. The main user interface for the web tool is shown in Figs. 4.10, 4.11, and 4.12.

4.5 IR-only taxonomy

After defining each class using visible and near-infrared data, the next step was to create a system of classification using only near-infrared data. The goals here are twofold: first, it allows classification of data sets in the near-infrared where important mineralogical information lies, second, it acts as a test for the limits of information we can extract with this incomplete wavelength range. We discuss the limits of incomplete data ranges further in Section 4.6.

For many objects, data exist in either the visible or near-infrared wavelength ranges but not both. While taxonomies such as the Bus system (Bus, 1999; Bus and Binzel, 2002a) are available for visible data, no system has been widely accepted for assigning classes to data existing only in the near-infrared. We have adapted our present taxonomy to interpret spectral data available only in the near-infrared range. This adaptive taxonomy is not meant to determine a definite class, but instead is an intermediate tool to indicate classes. We especially note that several classes in section 4.3.5 are carried over unchanged from the Bus taxonomy and are based exclusively on features present at visible wavelengths. Assignment to these classes (Cg, Cgh, Xc, Xe, Xk) requires visible wavelength data, therefore objects in these classes cannot be recognized by near-infrared-only data. Further discussion is provided in Section 4.6.

To study the ability to classify objects having only near-infrared spectral data we took the same 371 objects used in the original taxonomy but included only data longward of 0.85 microns, again splining

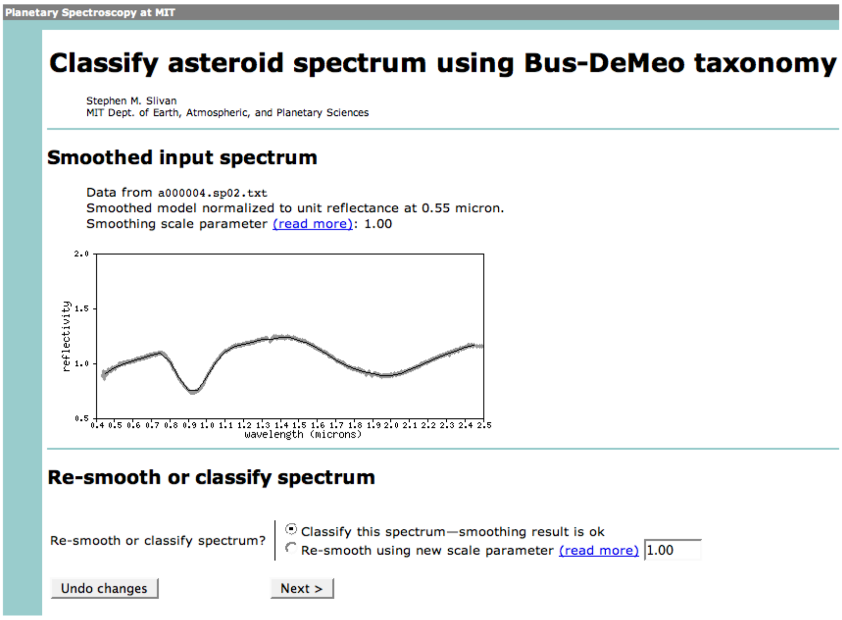

Figure 4.11: Taxonomy Web Tool Application Step 2: Smooth Spectrum. The user may increase or decrease the smoothing parameter (default = 1) to change the strength of smoothing. When the user is satisfied, s/he chooses “Classify this spectrum” and then “Next”.

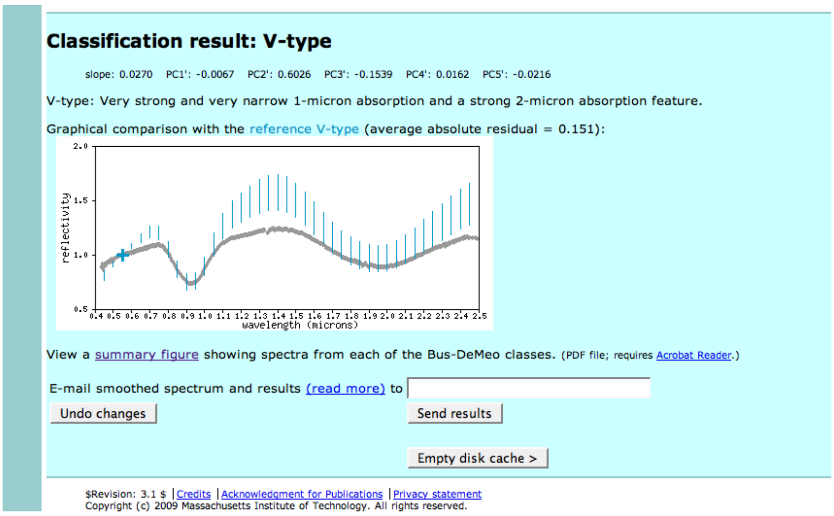

Figure 4.12: Taxonomy Web Tool Application Step 3: Classify Spectrum. Here the classification result is displayed along with principal component values and a plot of the spectrum compared to the average spectrum for that class. If more than one class is given, the user must inspect each plot and the average absolute residual of the input spectrum compared to the average for each class to determine the best classification for the object.

the data to smooth out noise. Our spline increments remained 0.05 μm covering the range of 0.85 to 2.45 microns resulting in 33 datapoints. We chose to normalize to unity at 1.2 microns, the closest splinefit wavelength value to 1.215 μm which is the isophotal wavelength for the J band based on the UKIRT filter set (Cohen et al., 1992). Next, we removed the slope from the data. As in the case with visible and near-infrared data we calculated the slope function without constraints, and then translate it in the y-direction to a value of unity at 1.2 microns. We then divide each spectrum by the slope function to remove the slope from the data set.

4.6 Limits of only visible or near-IR coverage

It is clear that both the visible and near-infrared wavelengths give important clues to the composition and alteration of asteroid surfaces, but what are the advantages of having both pieces of information? Does the visible wavelength range tell us everything we need to know? Can the near-infrared tell us everything we need to know?

4.6.1 Visible: The 1-micron band uncertainty

For S-complex and other olivine-rich asteroids, the visible wavelength is limiting because we cannot characterize the olivine and pyroxene content without the 1- and 2-micron bands. We find that the "very steep to extremely steep UV slope shortward of 0.75 μm, and a moderately deep absorption feature, longward of 0.75 μm" (Bus and Binzel, 2002a) that defines the Bus A-type surprisingly is not necessarily indicative of the very large 1-micron band seen in olivine-rich ($> 80\%$) spectra. Out of the 10 Bus A-types in our sample, half of them are not A-types in the Bus-DeMeo taxonomy (4 Sw-types, 1 L-type). An example of the divergence of Bus A-types in the near-infrared is shown in Fig. 4.13. When inspecting spectra classified as Bus Sa-types which are in the intermediate class between S and A because of the steep UV slope, we find that their near-infrared information excludes them from this status. All 12 Bus Sa-types within this sample were reclassified as Sw-, Sqw-, and S-types. While all of these objects had steep UV slopes, their 1-micron bands did not appear any more A-like than any other average S-complex spectrum. We find instead, that two objects previously classified as Sr-types under the Bus system do have the characteristic A-type wide and deep 1 micron absorption band, with very low overall slopes compared to A-types. Interestingly, these olivine-rich ($>\sim 80\%$) asteroids are hidden from identification in the visible-only wavelength range because their UV slopes are not nearly as steep and unique as Bus A-types.

Of the 16 Bus K-types in our sample, 5 of them were reclassified as L-types when greater spectral coverage is added. K-types differ slightly from L-types in the visible wavelength range, L-types having steeper UV slopes and a generally flatter spectrum past 0.75 μm, but in the near-infrared the K-types have a distinct 1 micron band, similar to an S-type although often slightly wider and more V-shaped rather than U-shaped, while L-types have much more subtle 1-micron bands.

We do find many strong consistencies between visible and near-ir data as well, which is very useful. Overall, we find that S-complex objects, remain S-complex with added near-ir data, even if the exact shape of the 1 micron band or depth of the 2 micron band cannot be entirely predicted by visible data. All three Bus Q-types in our sample remain unequivocally Q-types. All but one of the Bus V-types remain V-types. These classes are robust whether visible or the combination of visible and near-IR are used for classification.

4.6.2 Near-IR: S-complex and Q-types

The process of creating a decision tree to classify near-infrared-only asteroid data within the same system as the visible and near-IR data provided a rigorous test for the amount of information contained within the near-IR only data.

We started by performing Principal Component Analysis on the same 371 objects used in the original taxonomy, including only the data past 0.85 microns. The data were normalized to unity at 1.2 microns. We find that principal component space nicely separates featured from subtly featured classes in PCir2′ and PCir3′ space. Figure 4.14 shows this division in PC space. We learn that further subdivisions become

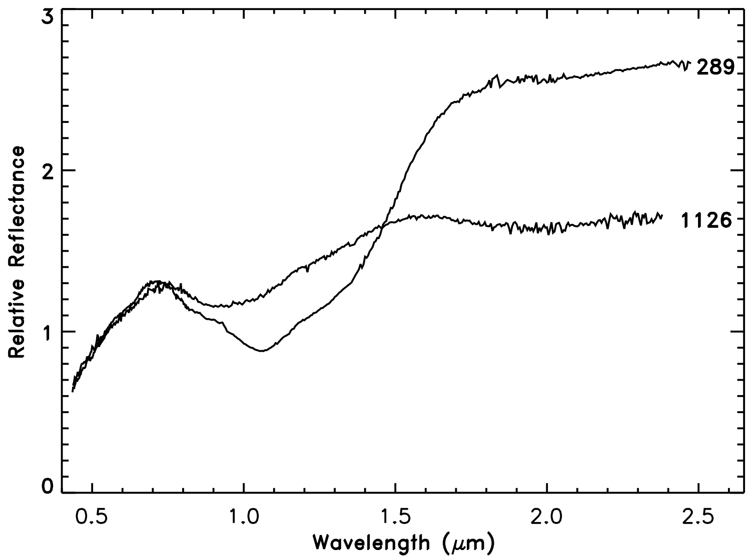

Figure 4.13: Shown here are two Bus A-types, asteroids 289 and 1126. While they have nearly identical behavior in the visible region, their spectra in the near-infrared region are significantly different. Asteroid 289 remains an A-type, but asteroid 1126 becomes and Sw-type.

much more difficult. PCir1′ and PCir2′ space is shown in Fig. 4.15. While some general boundaries can be constrained, the level of detail attained by using the visible plus near-infrared is not possible.

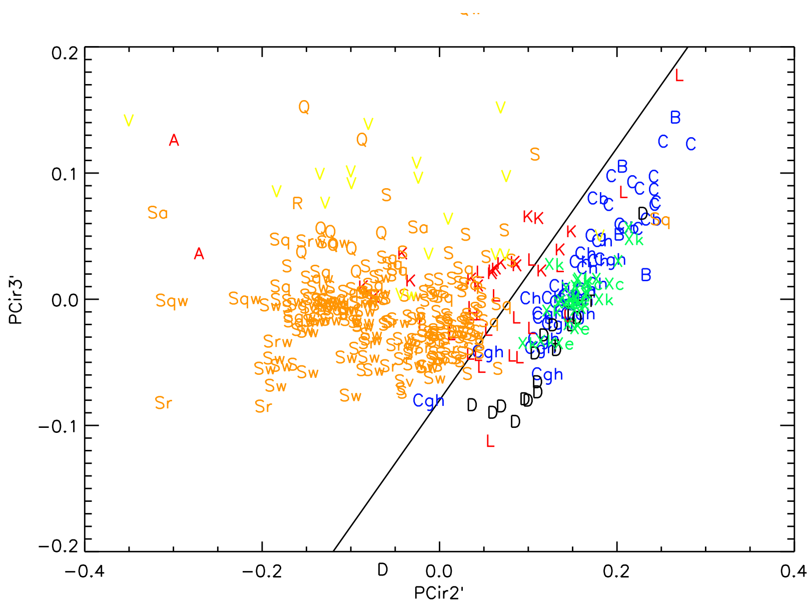

Figure 4.14: Plot of PCir2′ v. PCir3′. Here "featured" versus "subtly featured" objects are separated by the line.

Without the peak at the beginning of the 1-micron absorption feature, we lose important information about the depth and to a lesser extent the width of this feature making separation between S-, Sr-, Sq-, and

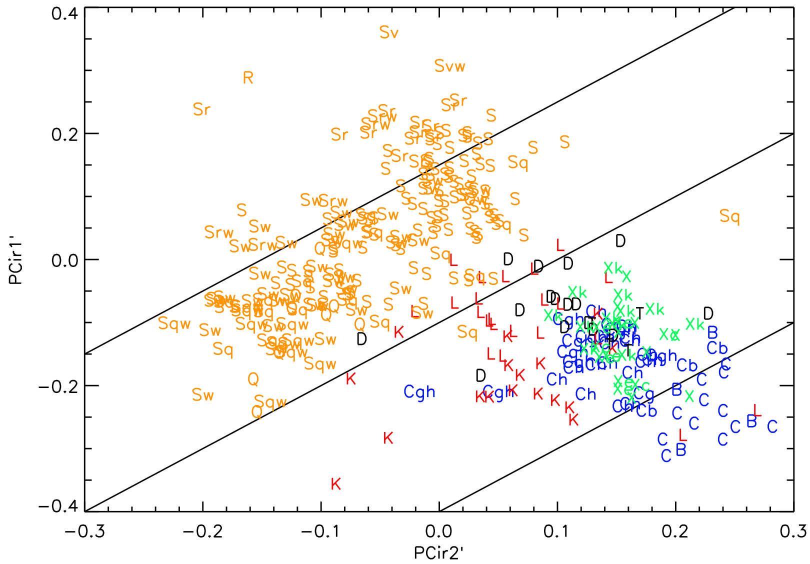

Figure 4.15: Plot of PCir1′ v. PCir2′. Here we create four segments separated by the three lines to help determine to which class an object belongs. Without the visible information, we cannot determine definite class boundaries in many cases.

even Q-types impossible. Principal component analysis cannot separate between these classes. Visual inspection of the overall band shape and width, however, usually allows a more definite classification than PCA could provide. For example, an Sv- and an Sq-type should never be confused even without the entire wavelength range, because Sv-types have extremely narrow bands, while Sq-types are much broader. Distinguishing between the S and Sq classes is often not clear with visual inspection because the differences between their bands are not always so great. We subsequently attempted to find a simple measure of band width to quantitatively separate classes.

By comparing the average spectrum of the S, Sr, Sq, Sv, and Q classes we find that the band minimum for Q is located at 1.0 micron, while for the others it is at 0.9 microns. We created a simple test by comparing the difference between the normalized reflectance value at 0.95 and 0.90 μm (f(0.95)-f(0.90)). We expected that all objects with negative values would be Q-types, because the reflectance values are still decreasing between 0.90 and 0.95 and the minimum has not yet been reached. We found that in general this could separate between Q-types and the others, but not definitively because the average spectra on which we based this method do not represent the entire range of spectra within that class. There were still S- and Sq-types that had minima at longer wavelengths. We attempted many other tests by comparing values and ratios of reflectance values at different points throughout the band. No test or combination of tests could adequately separate among classes.

We also found that many high-sloped S-types (Sw-types) were indistinguishable from lower-sloped objects in the sample with near-IR-only data. As stated in DeMeo et al. (2009a), because the entire 1-micron absorption band is not sampled, some depth versus slope information is lost, making it difficult, if not impossible, to distinguish between a steeply sloped spectrum with a shallow 1-micron feature and a spectrum with a lower slope, but a deep 1-micron feature. Figure 4.16 plots all S-complex, Q-type and R-type spectra in the sample normalized at 0.55 microns (right) and 1.2 microns (left). It is clear that the depth and width of the 1-micron band can be determined because the left peak is present. Figure 4.17 shows only the near-infrared wavelength region. Classification of these spectra is more difficult.

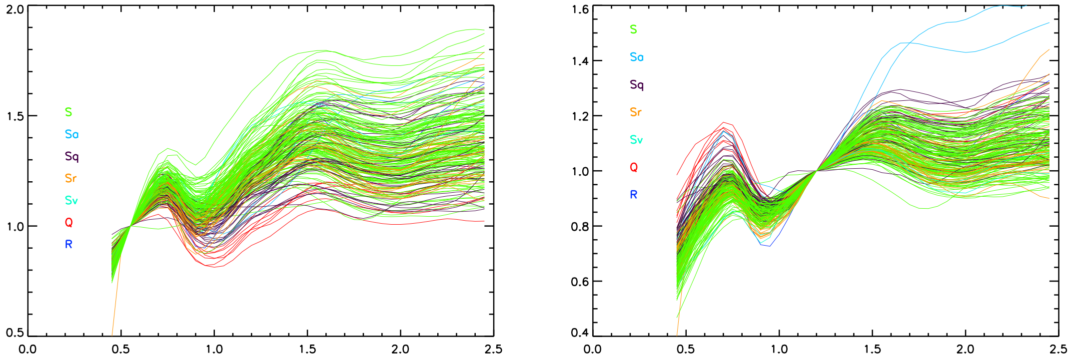

Figure 4.16: **Left:** Here we plot all S-complex, Q-type, and R-type spectra from our sample over the 0.45 - 2.45 micron wavelength range normalized to unity at 0.55 microns. The peak near 0.75 microns before the absorption band allows a clear determination of band depth, band width and overall slope of the spectrum. **Right:** Data here are normalized to unity at 1.2 microns to compare data over the entire range to near-IR-only data normalized at the same wavelength.

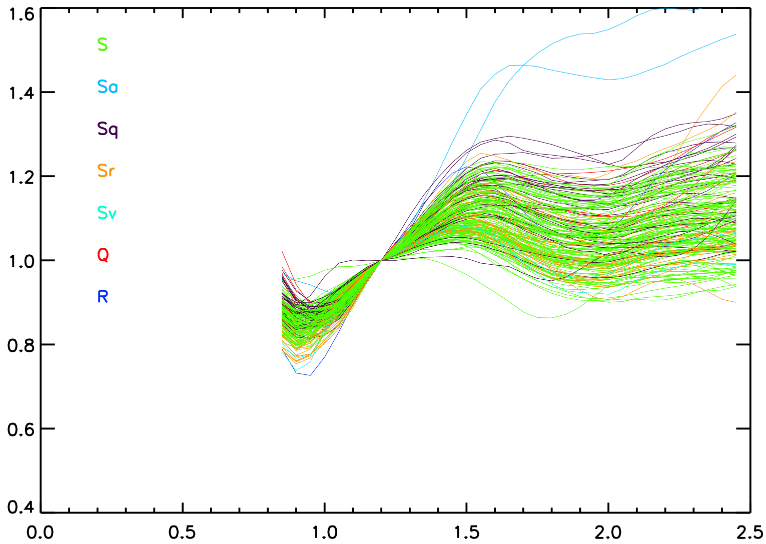

Figure 4.17: Here we plot all S-complex, Q-type, and R-type spectra from our sample over the 0.85 - 2.45 micron wavelength range normalized to unity at 1.2 microns. Without the peak near 0.75 microns before the absorption band, one cannot distinguish between a spectrum with a deep absorption and a low slope, or a shallow absorption and a high slope.

4.6.3 Near-IR: C- and X- complexes

While some information could still be salvaged about the 1-micron band by inspecting its shape and its width in the near-IR range, we learn about the degeneracy of subtly featured objects in the near-infrared. C- and X- complex objects do not have any strong distinguishing features in the near-infraread, with the exception of Xk-types which show a small feature near 1 micron, and C-types which tend to have a shallow, broad feature near 1-1.3 microns. Without the important visible wavelength information is it possible to separate these classes? Is there some slope information still retained? For example, do

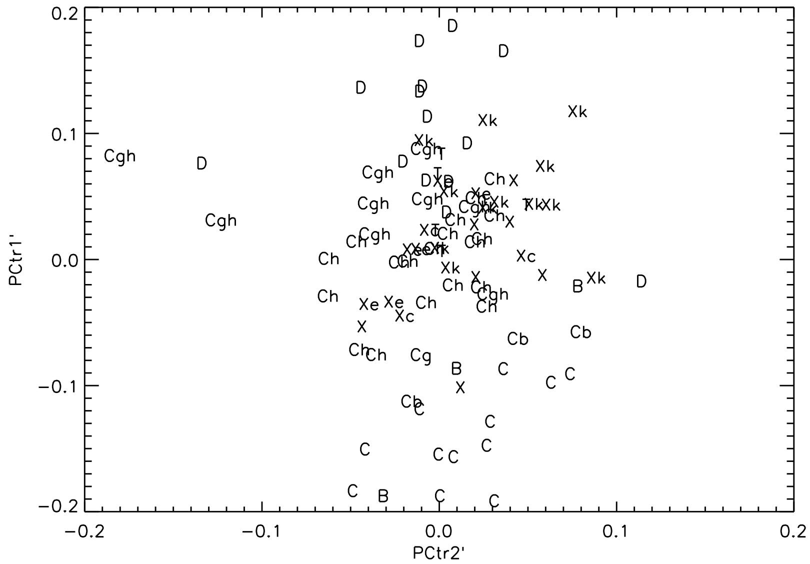

Figure 4.18: Plot of PC values from principal component analysis of *exclusively* C- and X-complex plus D- and T-type objects in near-infrared. Here the first two principal components are plotted.

C-complex objects have lower slopes than do X-complex objects in the near-infrared as is true in the visible?

We started with a comparison of slopes among classes. By comparing the average slope for each class and the range of slopes we find that there is a significant range of near-ir slopes within each class, much more than is seen in the well-contained visible wavelength region. There does not seem to be any clear slope boundaries, except for the exceptionally high-sloped D class.

Since no immediate answer could be found in our near-infrared PCA, we performed another principal component analysis on only subtly featured objects (including all C- and X-complex objects as well as D and T types). The advantage of isolating these objects is that PCA becomes more sensitive to subtle differences between these classes rather than the significant 1 and 2 micron features of S-complex objects.

We find by comparing the first principal component in this subtly-featured-only PCA that C-types tend to have the lowest PC values, while D-types tend to have the highest. These two classes, however, can already be separated by their slope for the D class and by visual confirmation of a subtle 1.3 micron feature for the C class. Unfortunately, no other clear separation between the B, Cb, Cg, Cgh, Ch, X, Xc, Xe, Xk, and T classes are evident. From this we can conclude that the C- and X- complexes can be defined exclusively with visible wavelength data and are entirely degenerate in the near-infrared. Figures 4.18, 4.19, and 4.20 plot the these objects in the various principal component spaces created in this analysis.

4.7 Albedo Distributions among Taxonomic Classes

The next step to refining taxonomy and ultimately our understanding of asteroid surfaces based on remote sensing is to include albedo information. Albedo tells us how dark or bright a surface is, helping constrain composition and age. It is generally found that the S-complex has higher albedos than the C-complex. Table 4.2 reports the average geometric albedo and the standard deviation for each class. The table also reports the debiased average albedos for NEOs calculated in Stuart and Binzel (2004), to compare to the primarily Main Belt sample in this taxonomy. Their objects were classified in the system of Bus and Binzel (2002a) and were grouped into complexes. The average C-type albedo for our sample is lower than for their C-complex. This is likely because we report separately the Cgh-types that have a higher average

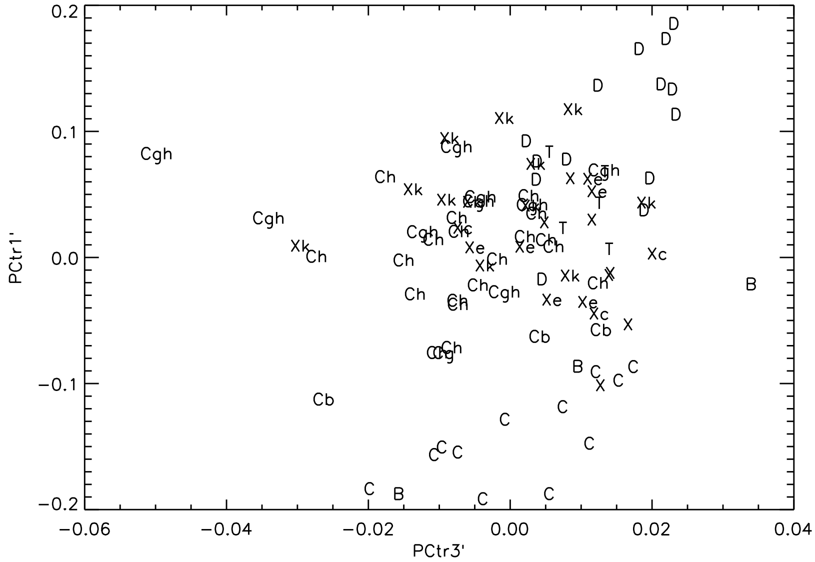

Figure 4.19: Plot of PC values from principal component analysis of *exclusively* C- and X-complex plus D- and T-type objects in near-infrared. Here the first and third principal components are plotted.

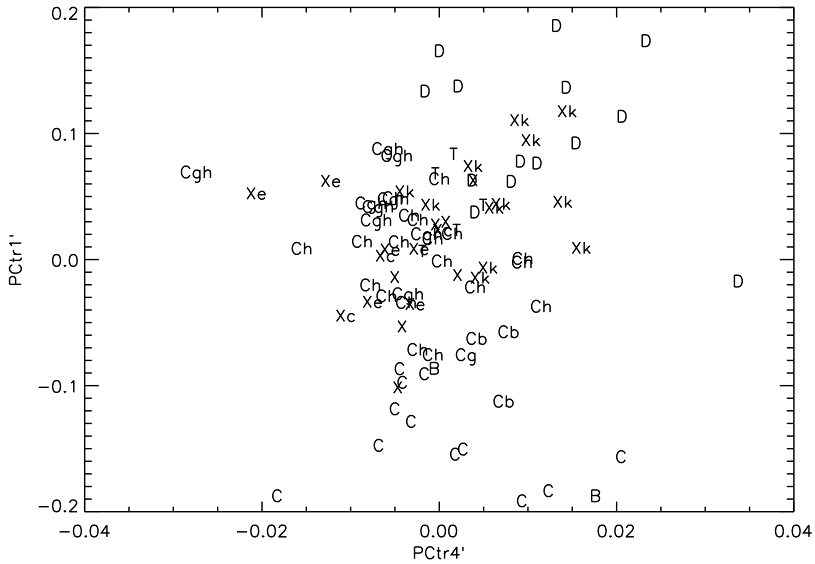

Figure 4.20: Plot of PC values from principal component analysis of *exclusively* C- and X-complex plus D- and T-type objects in near-infrared. Here the first and fourth principal components are plotted.

Table 4.2: Average Albedos for each Taxonomic Class

Class	Albedo[a]	Standard Dev	N[b]	NEO debiased albedo[c]
A	0.29	0.17	5	0.200±0.020
C	0.06	0.02	12	0.101±0.027
Cb	0.07	0.05	3	
Cgh	0.13	0.14	10	
Ch	0.06	0.01	16	
D	0.09	0.06	12	0.042±0.013
K	0.15	0.05	13	
L	0.15	0.05	14	
S	0.21	0.05	63	0.239±0.044
Sw	0.20	0.08	7	
V	0.27	0.14	3	0.417±0.147
X,T[d]	0.06	0.03	7	0.072±0.025
Xc	0.17	0.10	3	
Xe	0.14	0.03	3	
Xk	0.15	0.08	16	

[a]Average geometric albedo per class. Albedo values are from Tedesco et al. (2002)
[b]Number of objects used for average
[c]NEO debiased average albedos are from Stuart and Binzel (2004). They group the classes differently then we have here.
[d]Because the X and T classes are spectrally very similar we combine them for albedo averages to increase the sample size.

albedo. The average NEO albedo for D-types is lower than this sample, however it is based only on one albedo measurement. The V-types also have very different average albedos, although each of our samples only contains 3 objects. Vesta has a high albedo (0.42) matching those of the NEOs, which presumably came from Vesta. The two other Main Belt V-types with albedo measurements (1904 Massevitch and 1459 Magnya) are not in the Vesta family.

Figure 4.21 plots the distribution of geometric albedos for each taxonomic class for which there are at least three objects with albedo values. The albedo data is from Tedesco et al. (2002). While many classes have a small sample size, there are a few trends that can be noted. Many classes (C, Cb, Ch, S, X, T, and Xc) have reasonably well constrained albedos that vary by less than 0.15. The S-type has a broader overall range because of its large sample size, but the large majority of objects do not vary more than the other classes. Some classes (K, L, Sq, and Xk) have wide distribution of albedos among objects. Other classes have "albedo anomalies" in their class, with just one or two outliers with a significantly high albedo. While most Cgh-types have albedos less than 0.1, there is one object with an albedo greater than 0.4. Vesta is the high albedo outlier among V-types; however, the sample size is very small. Among D-types there are also two high albedo outliers. D-types are traditionally thought of as very dark, with very low albedos. The existence of two higher albedo objects forces us to recognize that this class of very red objects likely has a variety of different surfaces.

Larger samples, including objects classified under the Bus or Tholen systems, or simple classifications from SDSS colors, should provide much more insight on the albedo distributions of asteroid populations. The information provided by albedos of objects classified in this system, however, have the advantage of knowledge of spectral behavior in the near-infrared. With a larger sample, it would be worth reexamining the albedo distributions particularly among S-complex objects for which near-infrared data is critical.

4.8 Conclusion

An extended taxonomy was presented here using Principal Component Analysis and visible features to characterize visible and near-infrared wavelength spectra. The system, based on the Bus visible taxonomy from Bus (1999); Bus and Binzel (2002a), has 24 classes compared to 26 in the Bus system. We eliminated three classes: Ld, Sl, and Sk. All the Bus S subclasses (Sa, Sl, Sk, Sq, Sr) had objects that merged back

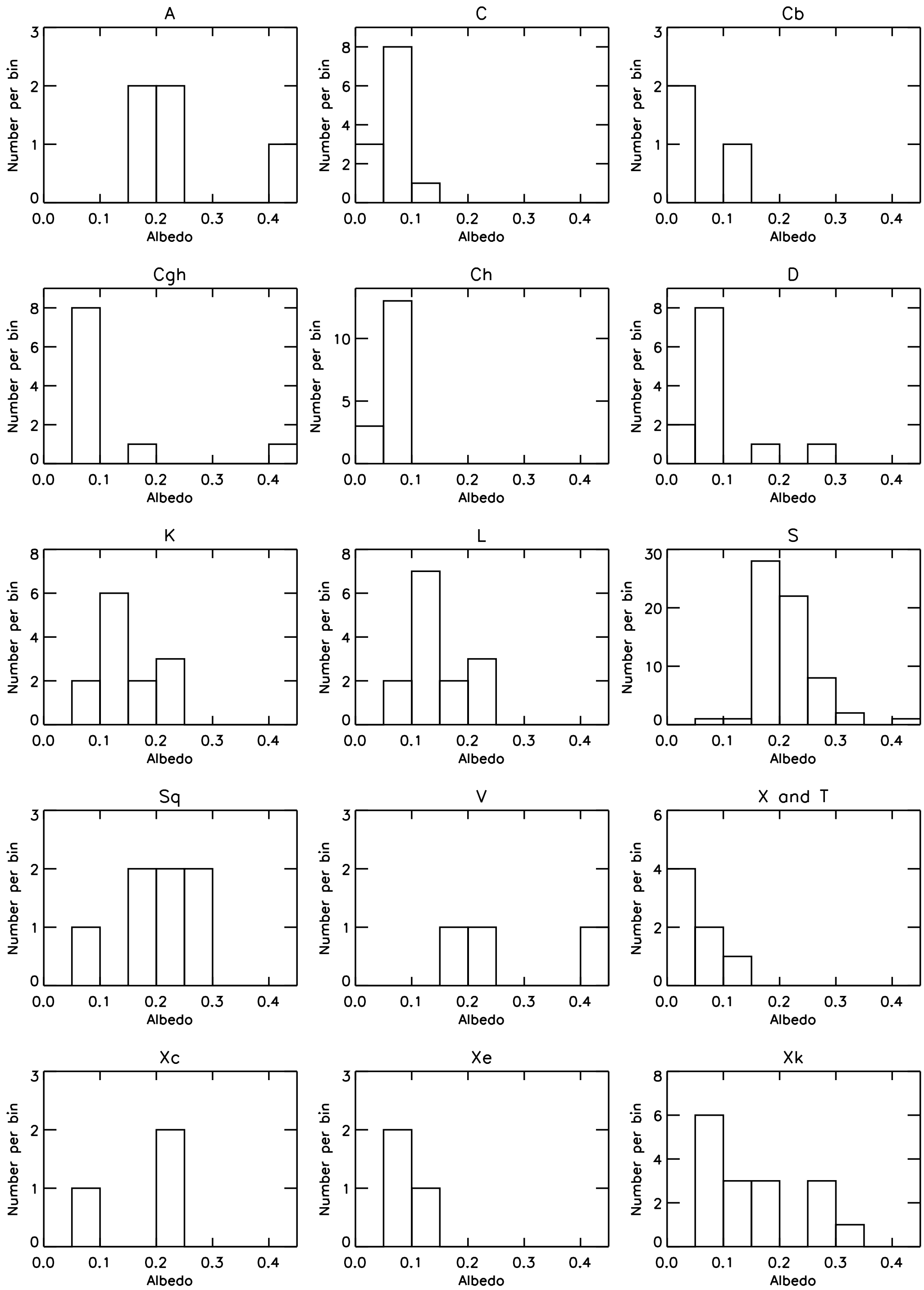

Figure 4.21: The distribution of object geometric albedos per class. Albedos are from Tedesco et al. (2002). Only classes with at least 3 objects with albedo measurements are presented. The bin width is 0.05, and the last bin represents any albedo value greater than 0.4. The mean and median albedo error for the sample are 0.02 and 0.01, respectively. The largest error by far is for a single S-type with a measured albedo of 0.43±0.15.

into the S-class, although many Sq objects remained Sq and two Sr objects were relabeled Sa. A new intermediate class, the Sv-class, was created as a link between the S- and V-classes. High-sloped S, Sq, Sr, Sv, and V objects were given a w notation to indicate possible weathering, but this notation does not constitute a new class. Many of the classes that lie left of line α in PC2$'$ versus PC1$'$ space are either featureless or exhibit only small features at visible wavelengths identified by Bus (1999); Bus and Binzel (2002a). It is still necessary to use these visible features to distinguish the classes because there are no other corresponding features at near-infrared wavelengths. We have also devised a method to categorize data when solely the near-infrared wavelength range is available, however, without visible wavelength information, the near-infrared taxonomy supplement cannot definitively classify many types especially those in the C- and X-complex, as many of those classes are defined only by visible wavelength features.

By attempting to create a supplementary method of classifying near-infrared-only data, we simultaneously provided a test for the limits of having only visible or only near-infrared wavelength data. We find that visible wavelength data is strongly indicative of near-infrared behavior. Spectra classified as A-types however, often do not have the expected strong 1-μm absorption band. We learn that the strength of the dip seen toward the end of the visible regime that indicates the presence of a 1-μm feature, does not necessarily indicate the shape and depth of that feature. This prompted the change of designation among some objects labeled L and K, as well as among some S-, Sq, and Sr- types.

Additionally, we have found the limits of PCA. When near-infrared data are added, PCA is overwhelmed by the 1- and 2- micron features and does not do an adequate job of identifying more subtle features such as those found in the visible region in the C- and X- complex spectra. Even though a more quantitative approach is preferred for distinguishing groups, being able to visually identify and define characteristic features of a spectrum has proven valuable.

The ultimate goal of a taxonomy is to lead toward a better understanding of the mineralogic composition of asteroids. While grouping asteroids is a useful tool, without mineralogic insight behind it, taxonomy is “just a letter”. Many methods of analysis are used to interpret the mineralogy of asteroid spectra. Band depths and widths are used to interpret olivine and pyroxene contents (e.g. Gaffey1993). Modeling, such as that described in section 3.2.1, are also used to estimate abundances by reproducing a spectrum using optical constants created from laboratory spectra. Meteorites, which are actual asteroid samples that can be measured in the lab, are also enormously important for understanding the composition of asteroids. Linking meteorite types with asteroid classes can provide important mineralogic constraints.

There is much future work to be accomplished for asteroid taxonomy. More complete systems in the future should include more than spectral information. Including visible albedo, radar albedo, densities, polarimetric measurements, longer wavelength data, and dynamical families could more thoroughly separate types of objects. At the same time, a taxonomy must always strive to be user-friendly, for if it is overcomplicated or not easily accessible and understandable to the community, it is not providing its intended service.