Supplementary Material for
Dimensionality's Blessing:
Clustering Images by Underlying Distribution

Wen-Yan Lin
Advanced Digital Sciences Center
linwenyan.daniel@gmail.com

Siying Liu
Institute of Infocomm Research
szewinglau@gmail.com

Jian-Huang Lai
Sun Yat-Sen University
stsljh@mail.sysu.edu.cn

Yasuyuki Matsushita
Osaka University
yasumat@ist.osaka-u.ac.jp

We provide visual illustration of our top 30 distribution-clusters on data-sets evaluated in section 5. Results are shown in Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, with 5 images displayed per cluster. Our clusters are not only consistent constituent objects but also consistent in their poses. Our technique works on a large variety of images ranging from buildings, people, cats and object types. We also display cluster by k-means with our second order distance in Fig. 6, Fig. 7, Fig. 8, Fig. 9, Fig. 10 and traditional k-means in Fig. 11, Fig. 12, Fig. 13, Fig. 14, Fig. 15. Observe that even with the latest features, k-means clustering does not provide results that are qualitatively comparable with our distribution-clustering.

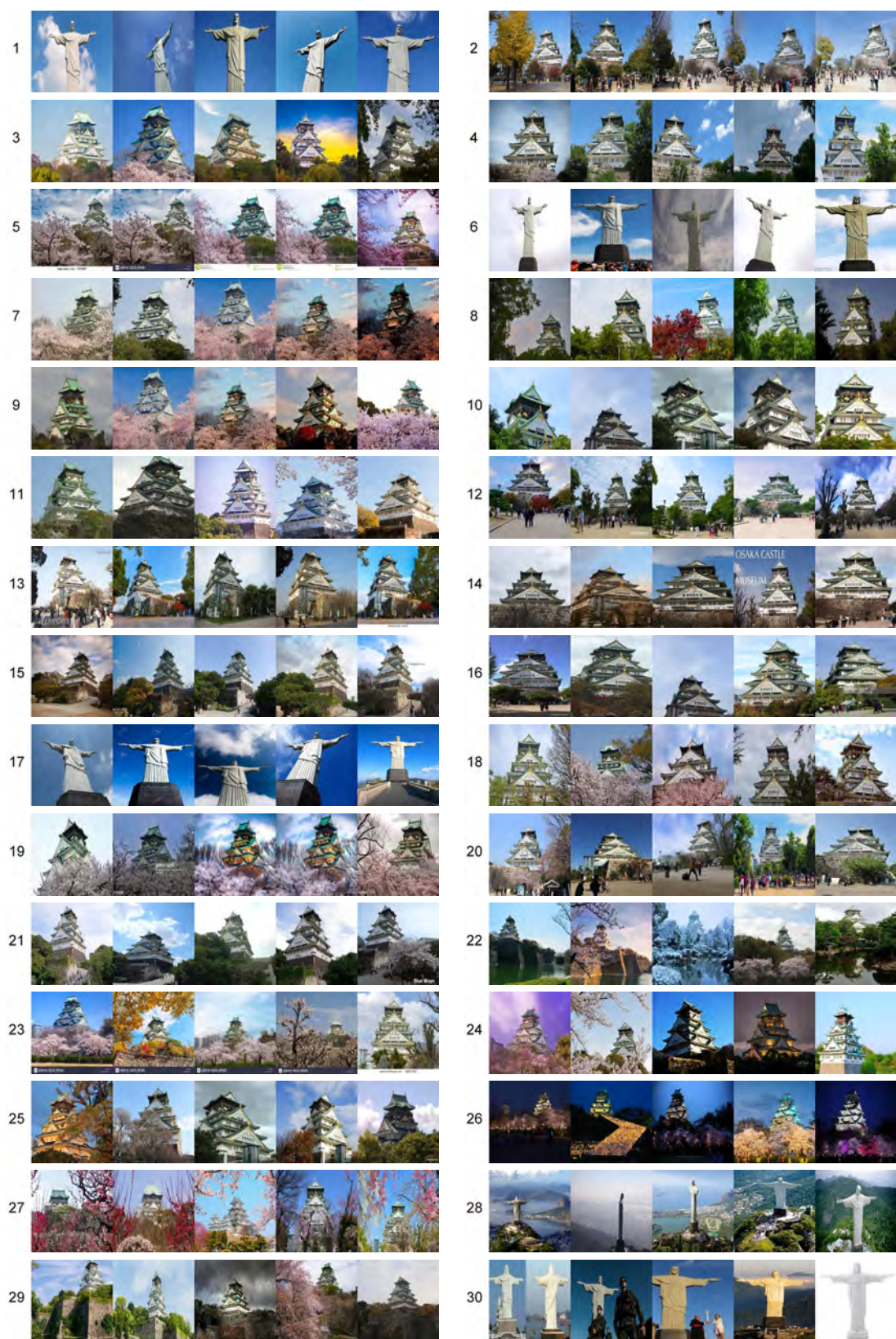Figure 1: Top 30 of our distribution-clusters from cats.

Figure 2: Top 30 of our distribution clusters from Internet.

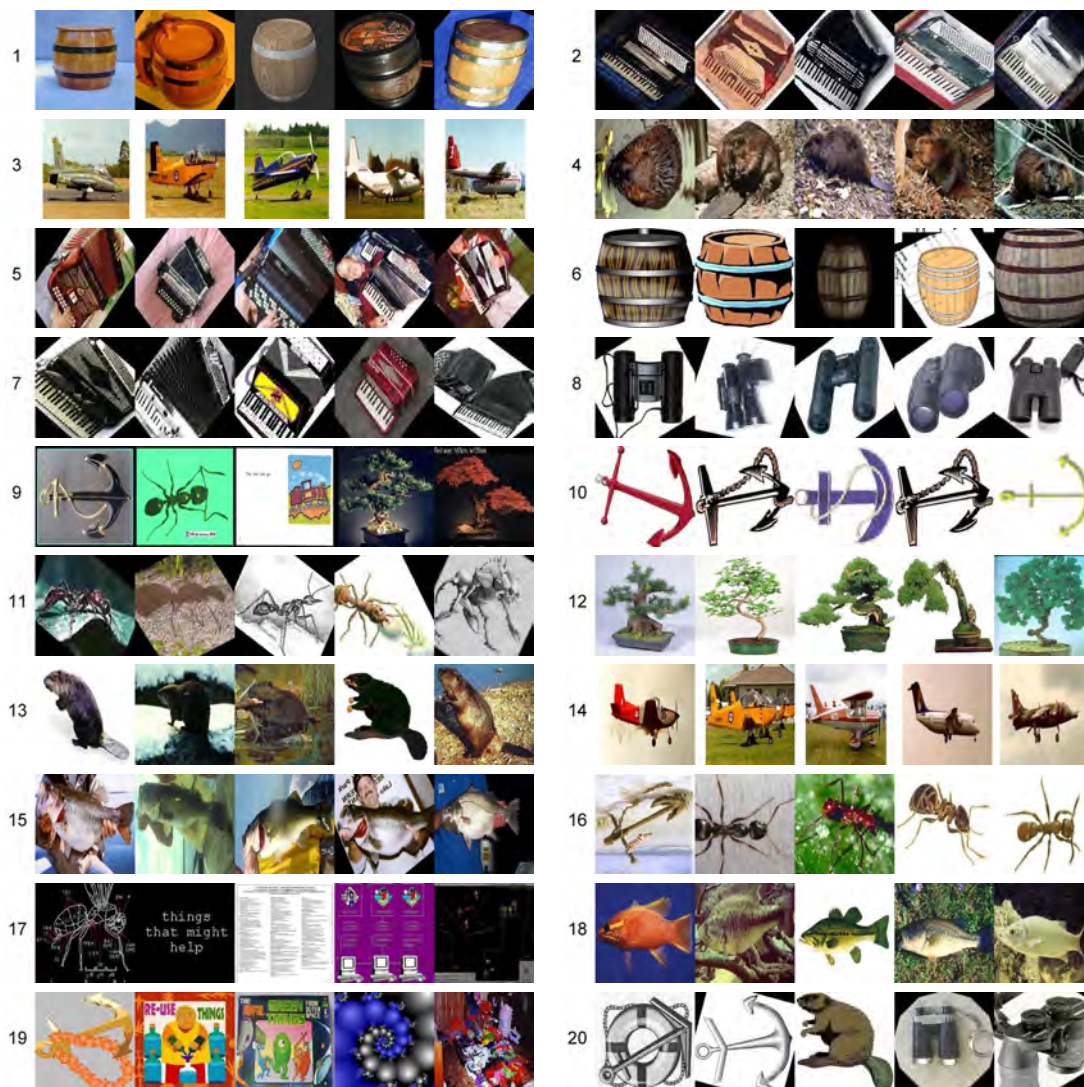Figure 3: Top 30 of our distribution-clusters from Mnist.

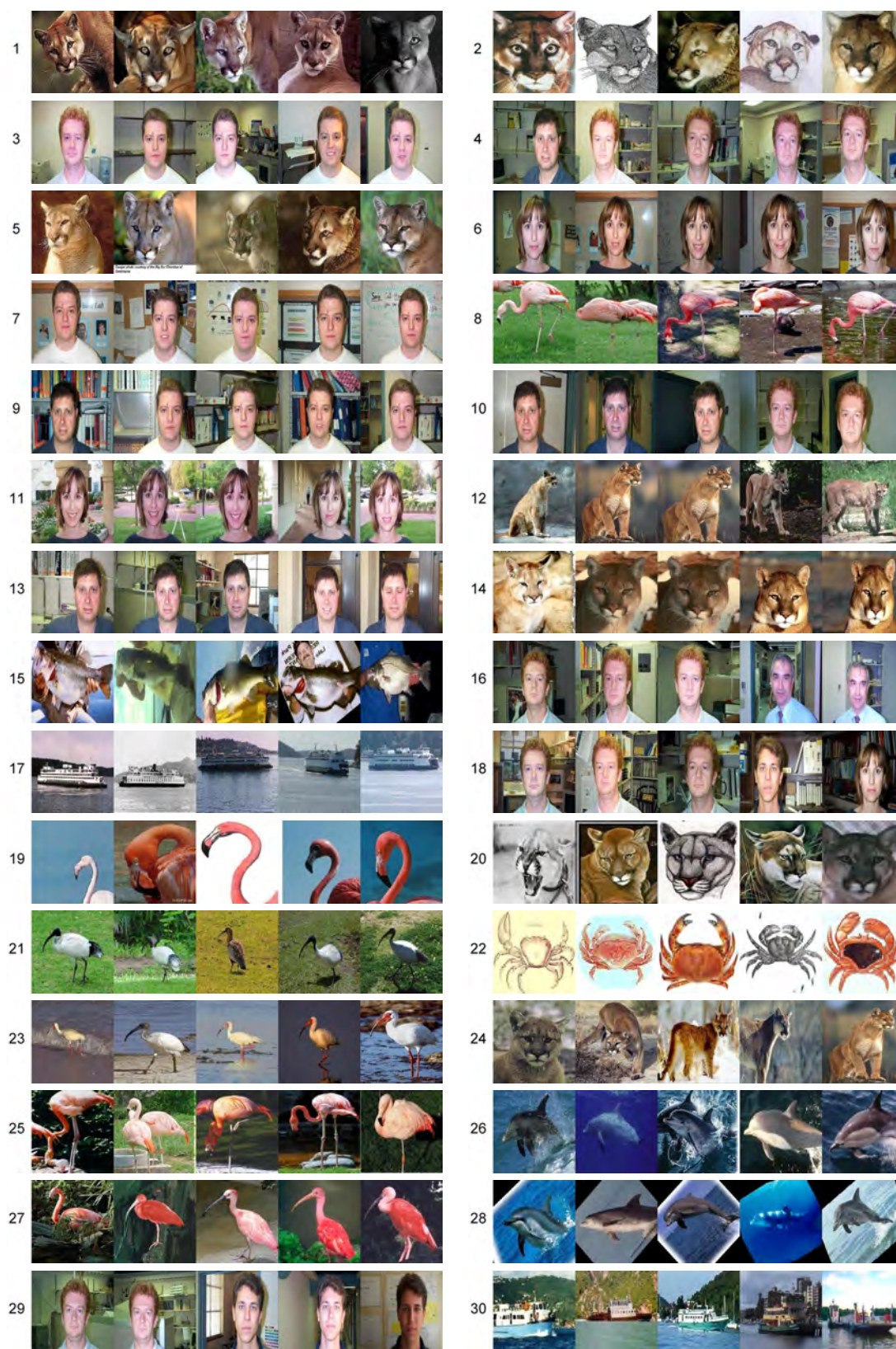Figure 4: All of our distribution-clusters from CalTech1.

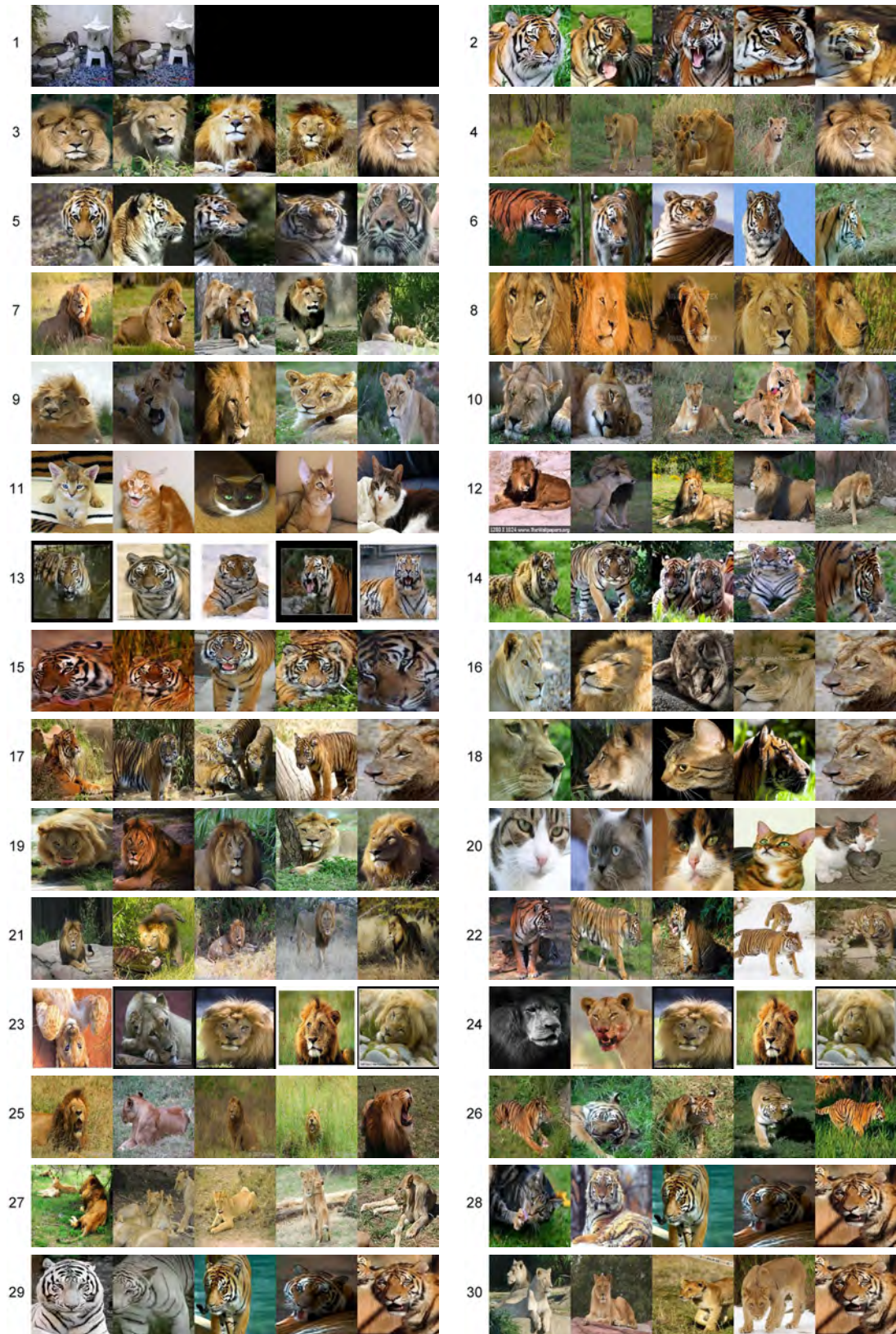Figure 5: Top 30 of our distribution-clusters from CalTech2.

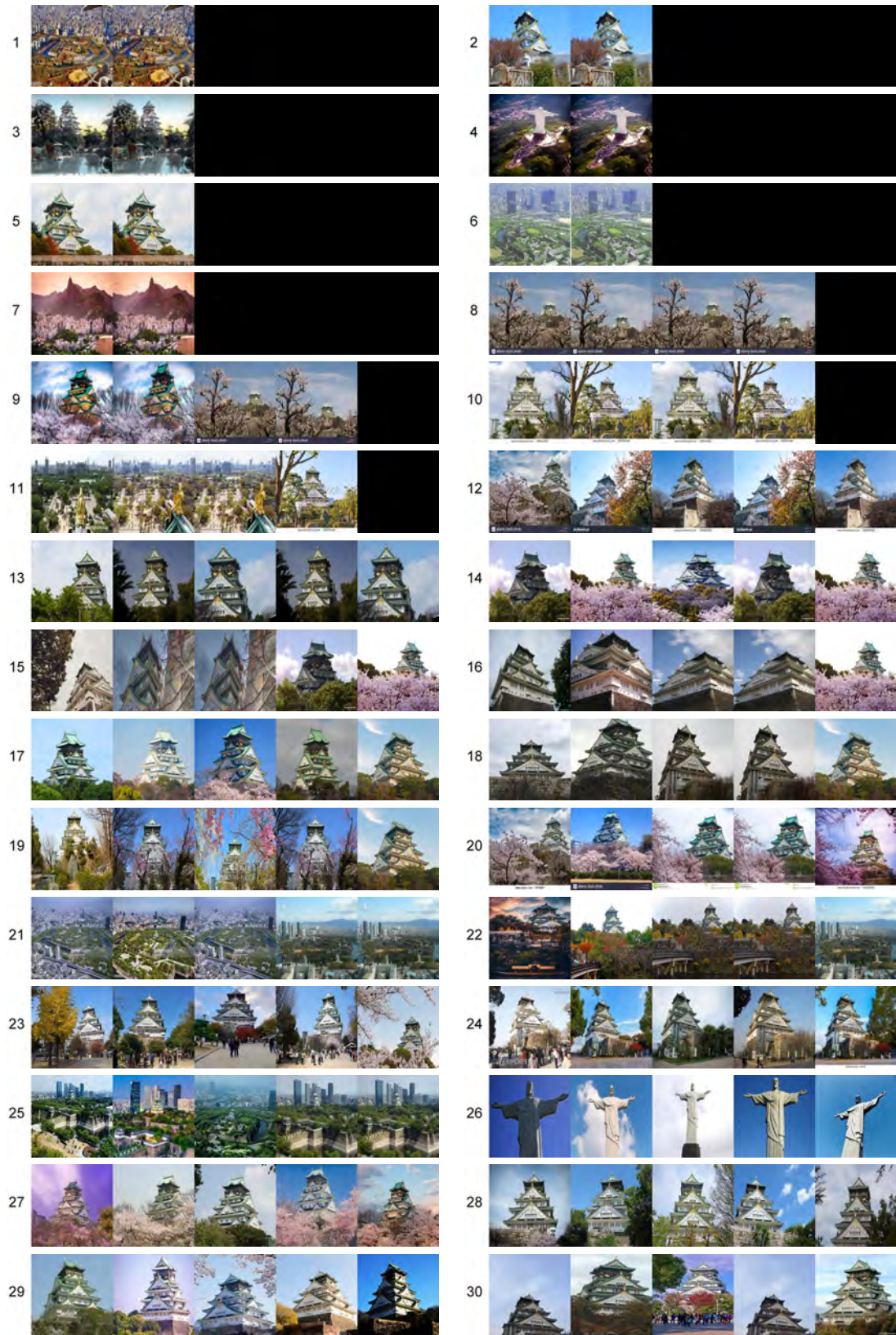Figure 6: Top 30 clusters, our second-order k-means, cats data-set

Figure 7: Top 30 clusters, our second-order k-means, Internet data-set
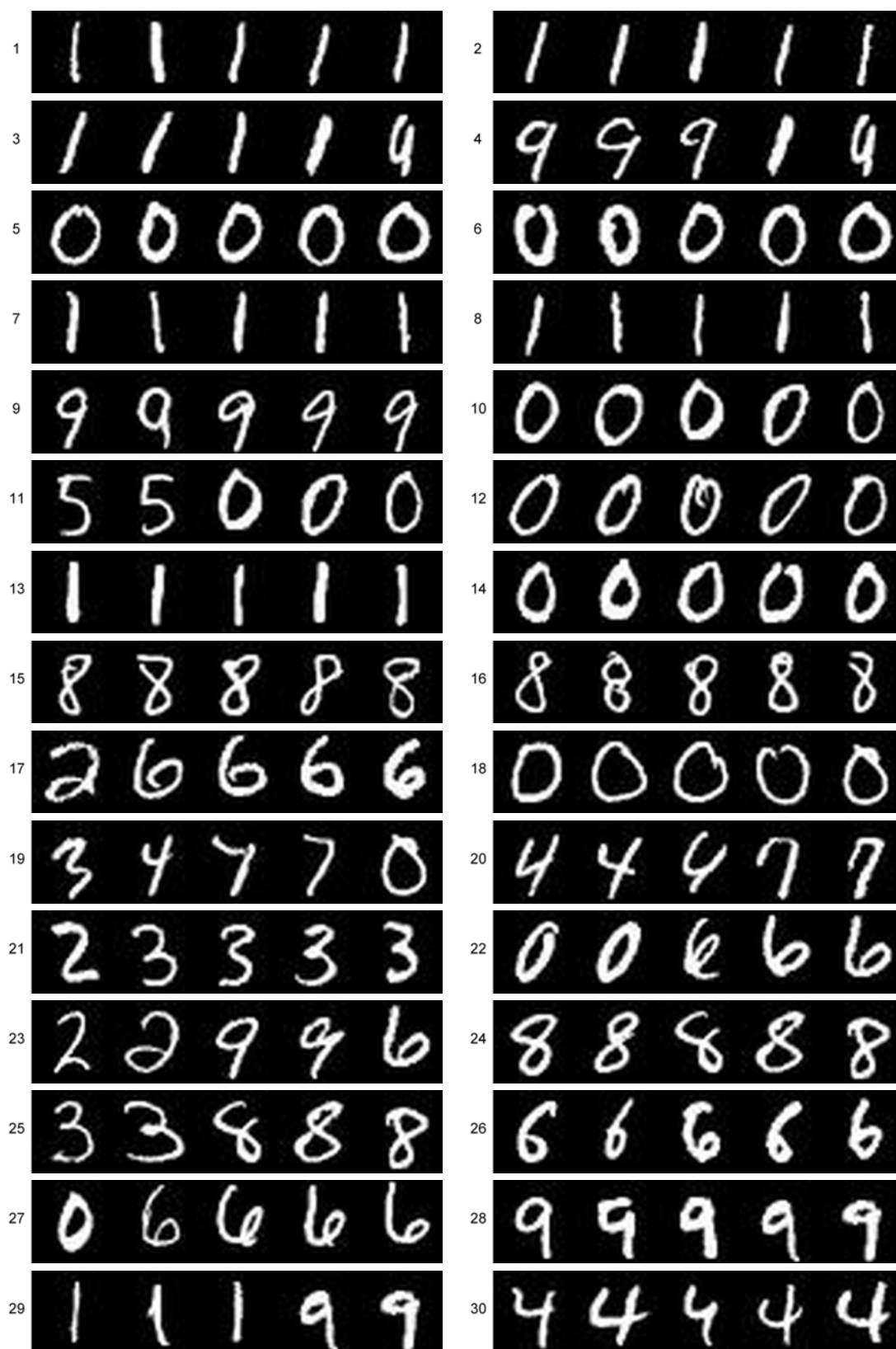
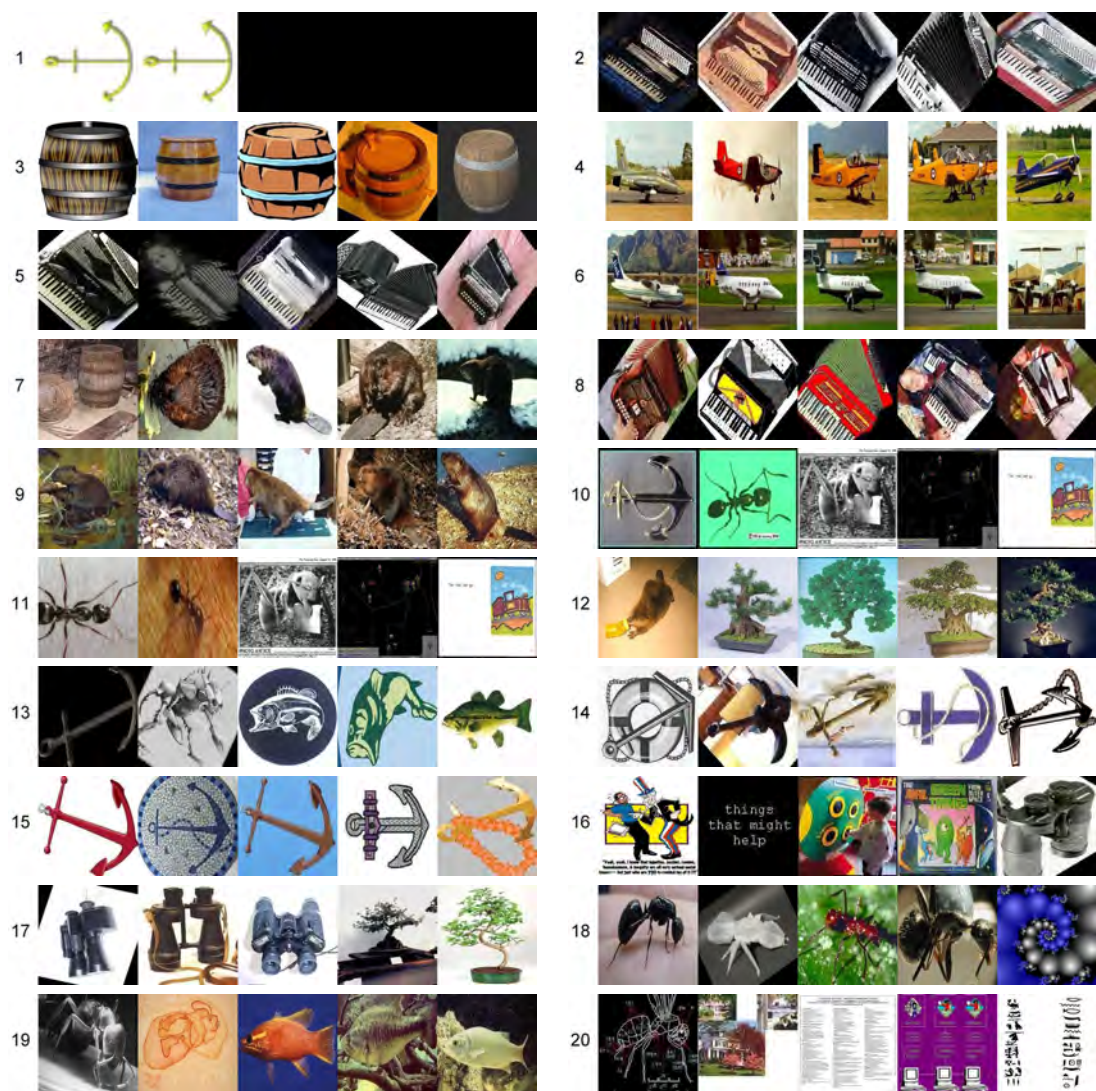Figure 8: Top 30 clusters, our second-order k-means, Mnist data-set

Figure 9: All clusters, our second-order k-means, CalTech1
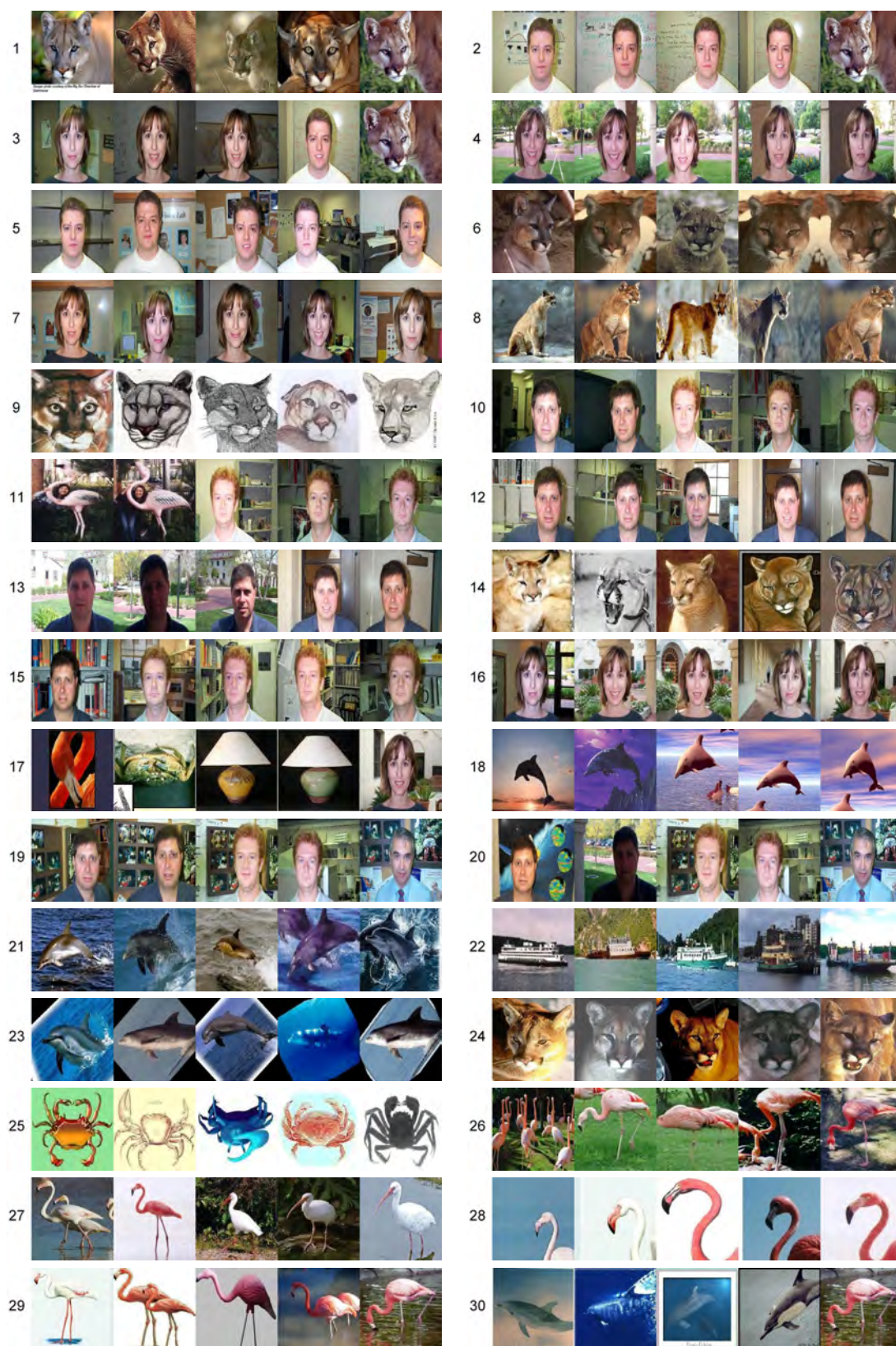
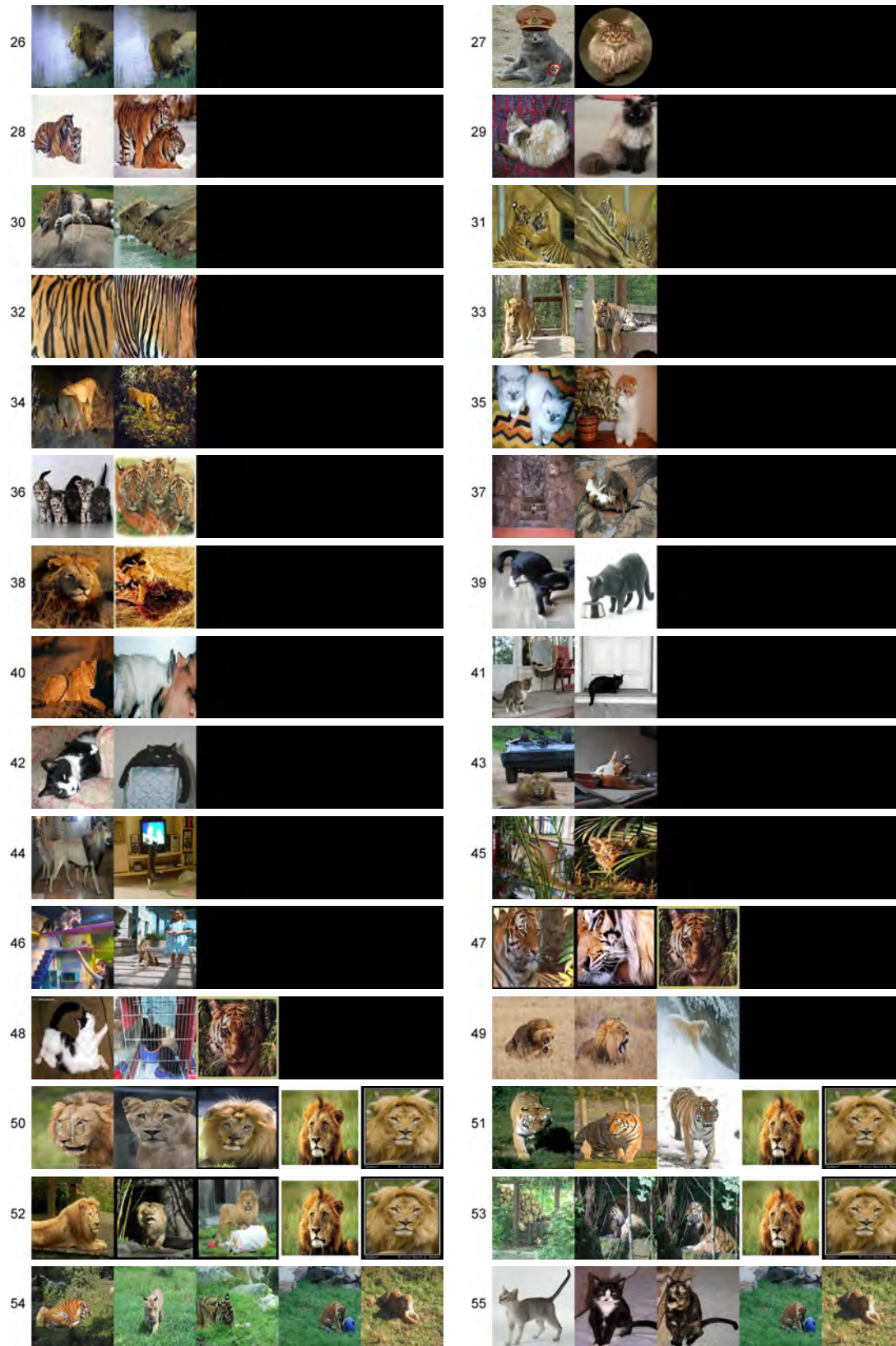Figure 10: Top 30 clusters, our second-order k-means, CalTech2

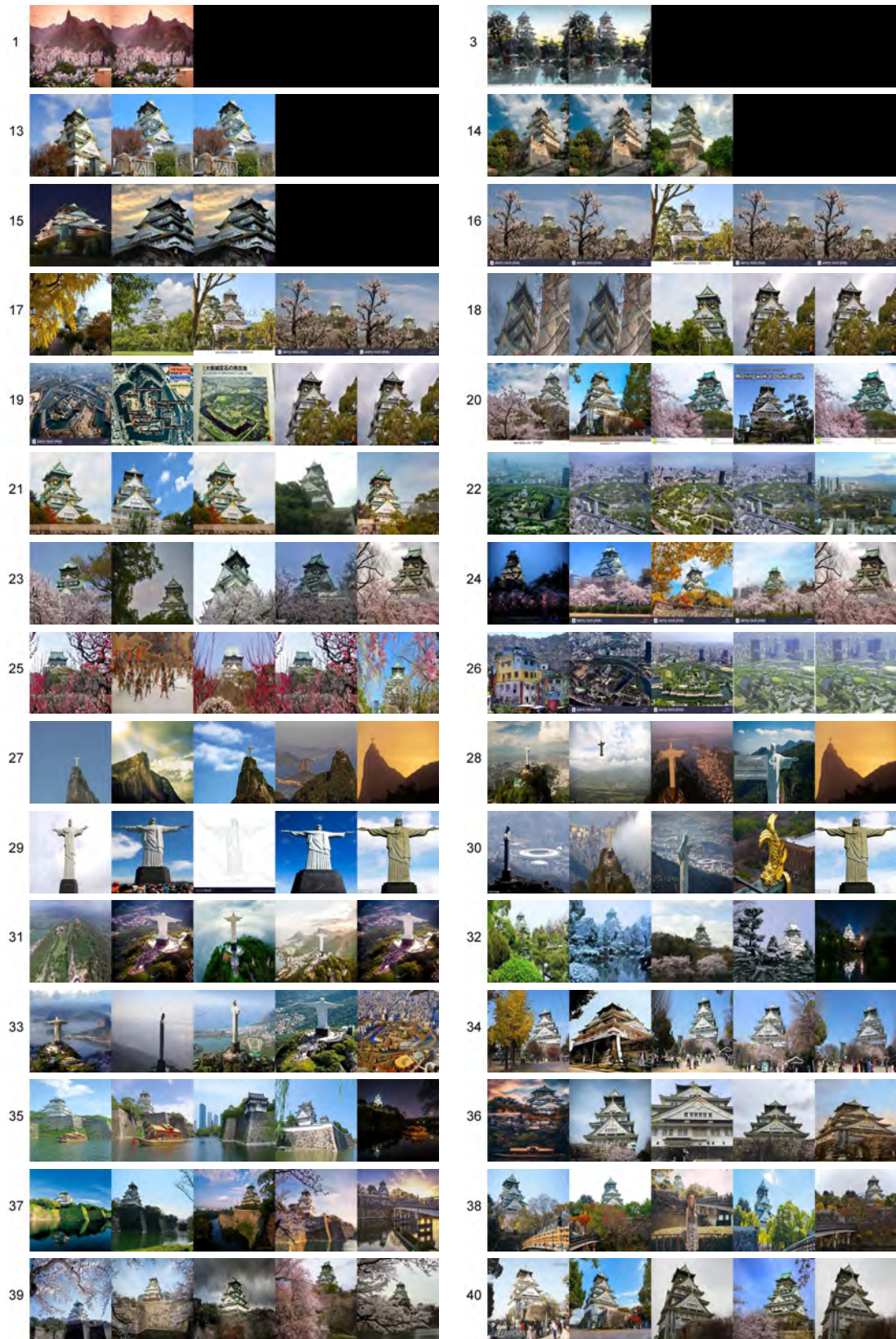Figure 11: Top 30 non-singleton k-means clusters from cats data-set.

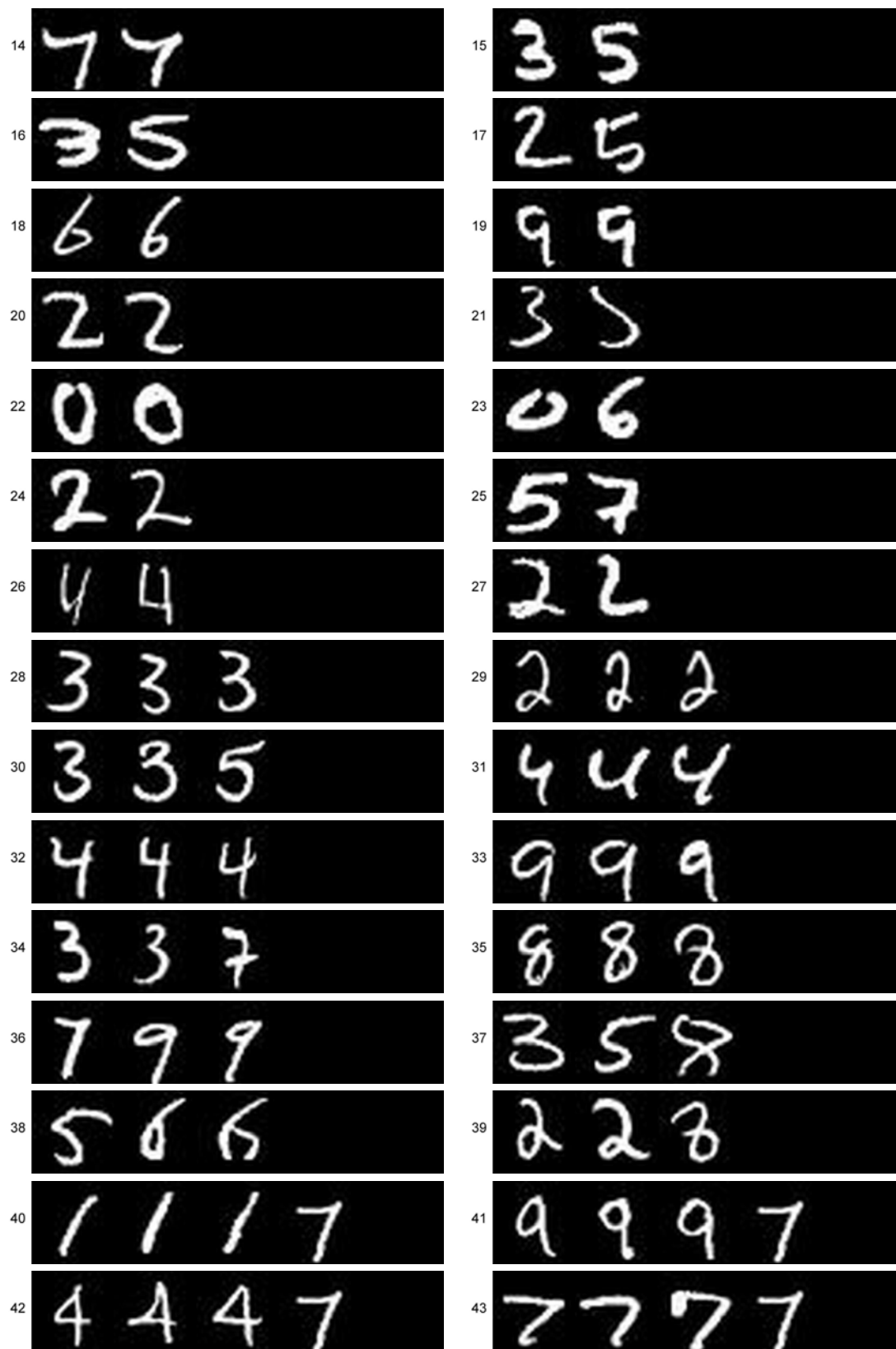Figure 12: Top 30 non-singleton k-means clusters from Internet data-set.

Figure 13: Top 30 non-singleton k-means clusters from Mnist data-set.

Figure 14: K-means clusters from CalTech1 data-set.

Figure 15: Top 30 non-singleton k-means clusters CalTech2 data-set.