

Online Navigation Summaries

Yogesh Girdhar and Gregory Dudek
{yogesh, dudek}@cim.mcgill.ca
Center for Intelligent Machines
McGill University

Abstract—Our objective is to find a small set of images that summarize a robot’s visual experience along a path. We present a novel on-line algorithm for this task. This algorithm is based on a new extension to the classical Secretaries Problem. We also present an extension to the idea of Bayesian Surprise, which we then use to measure the fitness of an image as a summary image.

I. INTRODUCTION

Navigation summaries are specialization of video summaries, where the focus is on video collected by a mobile robot, on a specific trajectory. We are interested in finding a few images that epitomize the visual information about the part of the world, which was traversed by the robot. Figure 1 shows a schematic example of a navigation summary.

This problem is related to the *Vacation Snapshot Problem*[1], where a tourist is trying to capture the most interesting observations of the journey, with a limited amount of film. A robot surveying an area might be taking continuous video as it is moving around, but the human monitoring the data might only be interested in a few images, which summarize what robot has seen; however, deciding which images are surprising or interesting, and epitomize the visual appearance of the world is highly context dependent.

A. Statistical versus Semantic Reasoning

Navigation summaries can broadly be classified into two categories, based on how the decisions to choose summary images are made. Images can be selected based on semantic cues regarding scene content, prior knowledge of the observer, and the exact context of the problem at hand; or by purely statistical reasoning, where the focus is on the information contained in the image, and previously selected images. The latter approach is, however, more tractable.

In this paper we will focus on the task of compiling navigation summaries based on statistical and information theoretic tools.

B. Online versus Offline

Navigation summaries can either be made offline, once the path traversal is over, or online. By online we mean that the decision to include an image in the summary set is made irrevocably, and immediately after it is acquired. For a mobile robot collecting thousands of images continuously, an offline algorithm comparing each image with every other image can be prohibitively expensive. Apart from computational cost, the ability to identify an image as being part of the

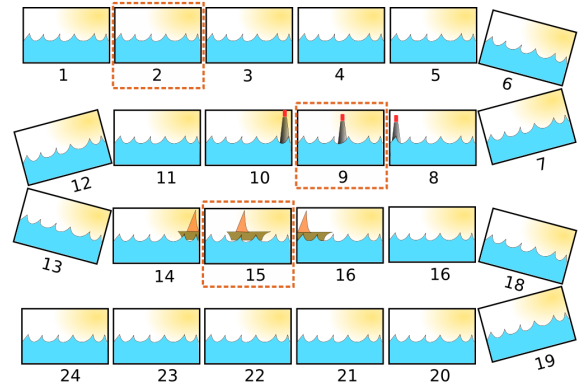


Fig. 1. An illustrated example of a navigation summary. The sequence of images represent the observations made by a robot as it is traversing a terrain. The dotted boxes indicates one possible choice of the summary images.

summary has several applications. For example, consider the task where we would like to drop a sensor node and take additional measurements at characteristic regions of a terrain. In that case, assuming each summary image represents a different region of the terrain, the robot could drop a sensor whenever a decision to include an image to the summary set is made.

Consider the special case of this problem where we only want one summary image, or want to drop only one sensor node. We can formalize and abstract this as a game where you are presented with a finite sequence of random numbers from an unknown distribution and then the goal is to identify the maximum number in this sequence. Here the maximum number might correspond to the amount of information an image has about the terrain, and hence maximizing it will give us the most important image. This problem is known as the *Secretary Problem* [2]. In this formulation, the task is posed as one of making a choice to select and hire irrevocably, the best secretary, with the interview score of the candidates represented as a sequence of random numbers.

If we are allowed to go back to a previously rejected sample(i.e. secretary), then our algorithm is trivial. We can just go through all the samples, identify the maximum value, and then in the end go back and choose the sample with the maximum value. Hence, the main challenge in the problem arises due to the fact that the decision made is irrevocable.

In this paper we will present an extension to the secretary

problem, which can then be used to select not just one, but several images that form a summary set; on-line. We also present a scoring function based on generalization of Bayesian Surprise, used to select the summary images.

II. PREVIOUS WORK

A. The Secretary Hiring Problem

The secretary problem has a long and varied history. Due to its broad relevance to different domains, it has been considered by different authors in several different contexts. It is generally accepted that Dynkin [2] was the first one to solve the problem formally. For an interesting discussion on the origins of this problem see [3]. Here is a description of the problem in its simplest form with the number of secretaries $k = 1$:

- You are given the task of hiring the best possible candidate for the job of a secretary.
- There are n applicants who have signed up for the interview.
- After interviewing each candidate you can rank them relative to all other candidates seen so far.
- You must either hire or reject the candidate immediately after the interview.
- You are not allowed to go back and hire a previously rejected candidate.

A typical strategy would be to just observe the first r candidates without accepting any, then find the highest score among them, and then hire the first candidate with score higher than that. This is known to be the provably optimal strategy for this problem. The problem now is to select the best value for r .

Let $\Phi(r)$ be the probability of successfully finding the highest scoring candidate, when we set the training interval to be r . We then have:

$$\Phi(r) = \sum_{i=r+1}^n P(S_i), \quad (1)$$

where S_i is the event that the i th candidate is the highest scoring candidate, and that our algorithm did not select any of the previous candidates. Hence we have:

$$\Phi(r) = \sum_{i=r+1}^n \frac{1}{n} \cdot \frac{r}{i-1} \quad (2)$$

$$\approx \frac{r}{n} \int_r^n \frac{1}{i} dx \quad (3)$$

$$= \frac{r}{n} (\log n - \log r). \quad (4)$$

Now to optimize $\Phi(r)$, we set the derivative equal to 0:

$$\frac{d}{dx} \Phi(r) = \frac{\log n - \log r}{n} - \frac{1}{n} = 0 \quad (5)$$

$$\implies r = \frac{n}{e}. \quad (6)$$

Here $1/n$ is the probability that the i th candidate is the highest scoring one, and $r/(i-1)$ is the probability none of the previous candidates were selected.

B. The Multiple Choice Secretary Problem

For the case when the number of positions which need to be filled is more than one ($k > 1$), there are several possible ways in which the above single secretary solution can be generalized.

Kleinberg [4] suggested an algorithm to maximize the expected sum of the scores of the candidates. The algorithm works by splitting the candidates into two roughly equal intervals, where the boundary is chosen randomly using a binomial distribution $B(n, 1/2)$. We then recursively apply the classic ($k = 1$) secretary algorithm to the first half of the candidates, choosing $l = \lfloor k/2 \rfloor$ candidates. While doing this we also find the l th highest scoring candidate from the first half and use this as a fixed threshold to select the remaining candidates in the second half.

Babaioff et al. [5], [6] suggest a simpler algorithm with the same goal of maximizing the expected sum of the scores of the selected candidates. A sliding threshold to choose the candidates. Algorithm 1 describes this approach.

```

Find the top  $k$  scores in the first  $r = \lfloor n/e \rfloor$  candidates,
without selecting any. Call this list of thresholds,
 $T = \{t_1, \dots, t_k\}$ .
foreach remaining candidate  $(x_{r+1}, \dots, x_n)$  do
  if candidate has score higher than the minimum
  score in  $T$  then
    Hire the candidate.
    Remove the minimum value from the set  $T$ .
  end
if  $T$  is empty then
  break
end
end

```

Algorithm 1: MaxExpectedSumScores($\{x_1, \dots, x_n\}, k$)

Note that both Klenberg and Babaioff algorithms, are optimal in maximizing the sum of candidate scores. However, they assumes that the scoring function does not depend on the what has already been selected. Hence using these algorithms to select our summary images will only provide an optimal expected case solution when all images are statistically independent and do not share any information. This is typically not true. For example, images adjacent in time are bound to be quite similar. In Section III-A, we will present a new algorithm which does not have this deficiency.

There are several other variants of this problem which exist in the literature. Freeman [9] reviews some of these variants.

C. Video Summaries

Most of the existing work related to the problem of navigation summaries is in fact on the more general problem of video summarization.

The most familiar and commonplace method for summarizing video is to subsample in temporal domain (i.e. viewing while fast forwarding). This approach works well for limited compression rates, but is rarely used for speeds above 32 times real-time. Moreover, the approach works well

for manually crafted (commercial) video where the length of time that a topic remains on-screen is generally proportional to its importance.

Smith and Kanade looked at producing video “skims” by depicting statistical summaries of video content accompanied by selected audio extracts [10]. That work employed the TFIDF (Term Frequency/Inverse Document Frequency) metaphor from text document indexing to index the audio data in a video stream. They used color histograms for video segmentation in the temporal domain. In that work, subsystems were also used to detect text captions and human frontal head views, in order to place emphasis on content that was particularly important in the context of summarizing television footage.

In related work, Ngo, Ma and Zhang cluster video frames based on an motion model and then used normalized cuts to segment the resulting proximity graph, and represent each cluster by an exemplar [12]. The motion model was based on MPEG motion vectors. A skim was produced that reduce the video length by up to 90 *per cent* by retaining the frames that made the largest contribution to the successive changes in the image or audio content.

In the work of Gong and Liu [14], video skims are produced based on a selection of key-frames with high information content. The accomplished by projection into a subspace that computed with respect to color histograms of a set of 9 (3x3) sub-windows that cover each input frames. This estimate is produced by computing a singular value decomposition over video sequence to estimate both the average frame, and the distance of each from the Eigenspace defined by the SVD. The distinctiveness of a frame is then computed by measuring its distance from the mean of the average frame in the SVD subspace, and this provides a criterion for selecting key-frames. While acceptable video summaries are reported, the authors suggest that the technique has shortcomings when color information is not distinctive enough.

In [15], Ju et al. considered the task of summarizing videos of a specific context, in this case, videos of conference presentations. Takeuchi et al. in [16] produced video summaries using a set of *a-priori* manually classified images.

Related to this problem is the problem of clustering images that depict the same environmental structure, albeit from different vantage points [18]. That work directly compares all n^2 images in the video sequence with one another using wide-baseline stereo methods. While the results of such an approach appear somewhat promising, the computational cost for a full video sequence are prohibitive.

III. APPROACH

A. Algorithm Overview

Our objective is to find a small set of images that summarize a robot’s visual experience along a path. We are presented with a stream of images taken by the robot and after observing each image, we must make an irrevocable decision accepting or rejecting the image as a summary image.

We first consider the easier offline version of the problem, where we are allowed to choose our summary images after doing all the pairwise comparison of the images. If we assume that each selected image covers information about some fraction of the input images and the trajectory, then this problem is reduces to an instance of the *Set Cover* problem, which is known to be NP-hard [19]. Hence we are motivated to look for approximate solutions.

Algorithm 2 presents a greedy approach to this problem. Here k is the number of desired images in the summary set, \mathcal{X} is the set of input images, and \mathcal{S} is the current set of summary images. The scoring function $Score(X|\mathcal{S})$, computes a fitness score of the given image X , given the previously selected set of summary images \mathcal{S} . In Section IV, we will present an information theory based scoring function, optimizing the summaries for information gain. This algorithm has $O(N^2)$ computational complexity in number of images in the input set ($|\mathcal{X}|$).

```

if  $k = 0$  then
  return  $\mathcal{S}$ 
end
 $X_{\max} \leftarrow \operatorname{argmax}_{X \in \mathcal{X}} \operatorname{Score}(X|\mathcal{S})$ 
 $\mathcal{S} \leftarrow \mathcal{S} \cup \{X_{\max}\}$ 
 $\mathcal{X} \leftarrow \mathcal{X} \setminus \{X_{\max}\}$ 
return SummarizeOffline( $k - 1, \mathcal{X}, \mathcal{S}$ )

```

Algorithm 2: SummarizeOffline($k, \mathcal{X}, \mathcal{S}$)

Algorithm 3 is on-line algorithm, which approximates the results of the greedy off-line algorithm presented above. Here t is the current time, X_t is the latest image acquired by the robot, k is the approximate number of images we want in the summary, n is constant and represents the total number of images we are expecting to see in this run, and \mathcal{S} is the set of currently selected summary images. The algorithm first decides on an observation interval (t, t_{obs}) , where it find the threshold score $v_{\text{threshold}}$. After that, the first image which exceeds this threshold is chosen. It then recursively calls itself to process future images.

```

if  $t > n$  or  $k = 0$  then
  return  $\mathcal{S}$ 
end
 $t_{\text{obs}} \leftarrow t + r(k, n - t)$ 
 $v_{\text{threshold}} \leftarrow \max \{ \operatorname{Score}(X_t|\mathcal{S}), \dots, \operatorname{Score}(X_{t_{\text{obs}}|\mathcal{S}}) \}$ 
 $t \leftarrow t_{\text{obs}}$ 
repeat
   $t \leftarrow t + 1$ 
until  $\operatorname{Score}(X_t|\mathcal{S}) \geq v_{\text{threshold}}$  or  $t \geq n$ 
 $\mathcal{S} \leftarrow \mathcal{S} \cup \{X_t\}$ 
return SummarizeOnline( $t + 1, k - 1, n, \mathcal{S}$ )

```

Algorithm 3: SummarizeOnline(t, k, n, \mathcal{S})

Apart from the scoring function $Score(X|\mathcal{S})$, performance of this algorithm mainly depends on the observation

interval function $r(k, n)$, which decides on how many images to observe, before starting the selection process. We use analysis similar to the the classical Secretaries Hiring problem, to compute the optimal value for this observation interval.

B. Top k Secretaries Using Fixed Threshold

Our goal now is to compute the observation interval, which maximizes the probability of finding the k highest scoring images. We will present the approach as an extension to the classical secretaries problem discussed in Section II-A, where a single threshold is used to optimally select the top k highest scoring candidates. The threshold is set to be the maximum observed score in the first r candidates. We would like r to be a function of k , the number of secretaries we want, such that if k increases, then r decreases. We can compute optimal value for r by maximizing the probability of success $\Phi(r)$, where success is defined by the event that all of the top k highest scoring candidates have been selected.

Let J_i be the event that with the selection of the i th candidate, we have succeeded. We can then write :

$$\Phi^k(r) = P(\text{Success}) \quad (7)$$

$$= P(\cup_{i=1}^n J_i) \quad (8)$$

$$= P(\cup_{i=r+k}^n J_i) \quad (9)$$

We ignore the first r candidates since those candidates are never selected as per our algorithm definition, and then we can ignore the next $k - 1$ candidates since its impossible to select k candidates from $k - 1$ possibilities. Analogous to Equation 2, we can then write $\Phi^k(r)$ as:

$$\Phi^k(r) = \sum_{i=r+k}^n P(J_i) \quad (10)$$

$$= \sum_{i=r+k}^n \frac{k}{n} \cdot \frac{r}{i-1} \cdot \frac{\binom{i-r-1}{k-1}}{\binom{n}{k-1}} \quad (11)$$

$$= \frac{k}{n} \cdot \frac{r}{\binom{n}{k-1}} \cdot \sum_{i=r+k+1}^{n-1} \frac{\binom{i-r}{k-1}}{i}, \quad (12)$$

where $\binom{n}{k}$ is the binomial coefficient. Lets us examine the three components of Equation 11. The first term: k/n is the probability that the i th candidate is one of the top k candidates. The second term: $r/(i-1)$ is the probability that none of the previous candidates were the last of the top k selected candidates. These two terms are similar to the two terms in Equation 2. The third term: $\binom{i-r-1}{k-1}/\binom{n}{k-1}$ is the probability that all of the remaining $k - 1$ candidates have been selected. Combining these terms we get the probability of the event that we have successfully selected last of the top k candidates.

C. Optimal Training Interval

Let \hat{r} be the values of r for which the probability of success $\Phi^k(r)$ is maximum. We computed the \hat{r} for $n = 1000$,

and $k = 1 \cdots 20$. We approximate this in closed form as:

$$\hat{r}(k) \approx \frac{n}{ke^{1/k}} \quad (13)$$

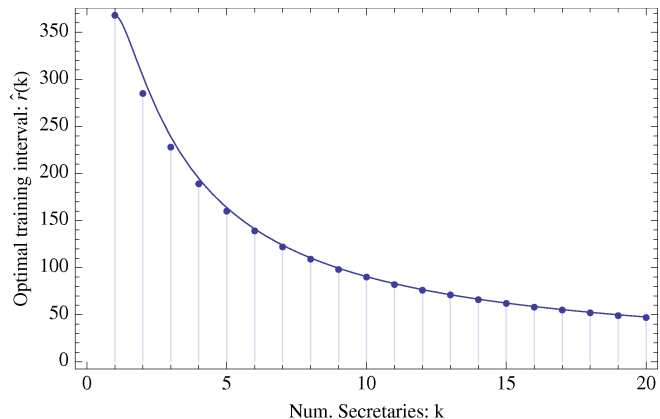


Fig. 2. Optimal training interval as a function of number of secretaries k , for $n = 1000$ secretaries. The dots represent the actual training interval, and we found that the function $\hat{r}(k) = n/ke^{1/k}$, represented by the curve above, is a good approximation of these points.

Figure 3 shows a plot of this approximation.

We can hence use this expression to set the observation interval in the algorithm described in Sec. 3. This will allow us to exploit this method in practice.

IV. SCORING FUNCTION

We would like our set of summary images to be chosen in such a way that they minimize the surprise in observing any of the other images.

A. Bayesian Surprise

Itti and Baldi [21] formally define surprise in terms of difference between posterior and prior beliefs about the world. Let our prior hypothesis H of the world be defined using the probability distribution $P(H)$. When a new data observation D is made, it can be called *surprising* if the posterior distribution $P(H|D)$ is significantly different from the prior distribution $P(H)$. One of the best ways to compare these distributions is using relative entropy or Kullback-Leibler(KL) divergence [22], which measure the information gain. We can then define surprise R as:

$$R(D, H) = d_{\text{KL}}(P(H|D)||P(H)) \quad (14)$$

B. Set Theoretic Surprise

Bayesian surprise measures the distance using KL divergence, between two distributions: $P(H|D)$ and $P(H)$. We propose modeling our prior and posterior using a set of distributions, where $P(H)$ is replaced by the set $\{P(H)\}_{H \in \mathcal{H}}$, and posterior $P(H|D)$ is replaced by the set $\{P(H^+)\}_{H^+ \in \mathcal{H} \cup \{D\}}$. We would now like to measure the distance between these two sets. Hausdorff metric provides a natural way to compute distance between two such sets. It defines this distance d as the maximum distance of a set to the nearest point in the other set:

$$d(A, B) = \max_{a \in A} \min_{b \in B} d(a, b). \quad (15)$$

Hence taking Hausdorff distance between our sets of posterior and prior hypothesis, where distance is KL divergence, we define *Set Theoretic Surprise* R^* as:

$$R^*(D, \mathcal{H}) = d(\mathcal{H} \cup \{D\}, \mathcal{H}) \quad (16)$$

$$= \max_{H^+ \in \mathcal{H} \cup D} \min_{H \in \mathcal{H}} d_{\text{KL}}(P(H^+) \| P(H)) \quad (17)$$

$$= \min_{H \in \mathcal{H}} d_{\text{KL}}(P(D) \| P(H)). \quad (18)$$

Applying this problem to our task of selecting summary images, we model the hypothesis using the set of selected images. Score of an image representing its suitability as a summary image, given a set of already selected summary images can then be defined as:

$$\text{Score}(X|\mathcal{S}) = \max_{S^+ \in \mathcal{S} \cup X} \min_{S \in \mathcal{S}} d_{\text{KL}}(P(S^+) \| P(S)) \quad (19)$$

$$= \min_{S \in \mathcal{S}} d_{\text{KL}}(P(X) \| P(S)), \quad (20)$$

where X is image whose score we would like to calculate, $P(\cdot)$ defines the distribution of image features, and S is an image from the set of selected images \mathcal{S} . In this paper we used a simple pixel color histogram as the image features.

V. RESULTS AND DISCUSSION

We tested our algorithm on several data sets, four of which are shown here. Each set used in this paper contains 64 images, and we want 8 samples from each set. Note that the offline algorithm we have described in this paper is guaranteed to give us 8 images. However, the online algorithm does not, because if in an iteration no images are found exceeding the threshold, then none are selected.

We tested our algorithm on three different types of terrains data.

Figure 5 show images from the street view data sets. We show the images selected by the online algorithm, and for comparison also show the 8 images returned by the offline algorithm. We see that 4/5 images in the online summary are part of the offline summary.

Figure 4 shows images from a simulated Martian analogue environment. We see that both the online and offline algorithms capture the visual appearance of the terrain well.

Overall we find the selections made by our algorithms clearly capture the diversity of the image types in input sample set.

VI. CONCLUSION AND FUTURE WORK

In this paper we looked at the problem of automatic generation of navigation summaries. A navigation summary is a small set of images, which capture the visual experience of a mobile robot, as it traverses a path. Our contribution to solving this problem is focused in two areas. Firstly we present a new extension to the classical secretaries problem, and use it to formulate a new online algorithm to build navigation summaries. Secondly, we present a generalization to concept of Bayesian Surprise, and then use it to pick the summary set.

In future we hope to improve our scoring function by considering more features. Some of these features could be

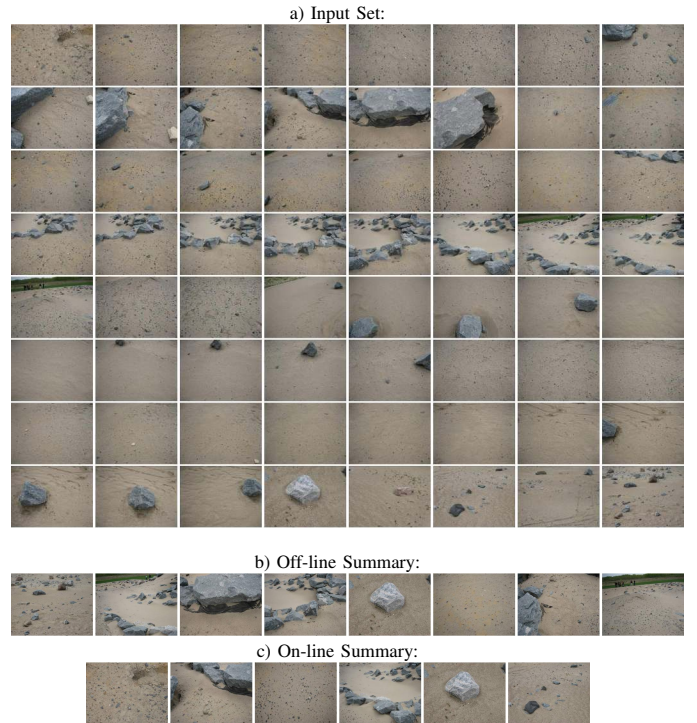


Fig. 3. Mars Dataset. a) 64 images acquired from the Mars Analogue Site operated by the Canadian Space Agency. b) Result of running the off-line algorithm. c) Result of running the on-line algorithm. We see good coverage of the variance in appearance of the terrain in both the off-line and on-line summaries.

explicitly defined, and be problem dependent; for example features like human faces. We also plan on looking at other statistical features which are problem independent; for example, frequency of different kind of textures in the images, or SIFT or SURF or MSER features.

VII. ACKNOWLEDGEMENTS

We would like to thank Google Inc. for generously allowing us to use their Street View data, and Canadian Space Agency for allowing us collect data at their Mars Analogue site. We would also like to thank Prof. Luc Devroye for his helpful guidance.

REFERENCES

- [1] E. Bourque and G. Dudek, "Automated image-based mapping," in *IEEE Computer Vision and Pattern Recognition (CVPR)—Workshop on Perception of Mobile Agents*, June 1998, pp. 61–70.
- [2] E. Dynkin, "The optimum choice of the instant for stopping a markov process," *Soviet Math. Dokl.*, vol. 4, 1963.
- [3] T. S. Ferguson, "Who solved the secretary problem?" *Statistical Science*, vol. 4, no. 3, pp. 282–296, 1989.
- [4] R. Kleinberg, "A multiple-choice secretary algorithm with applications to online auctions," in *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2005, pp. 630–631.
- [5] M. Babaioff, N. Immorlica, and R. Kleinberg, "Matroids, secretary problems, and online mechanisms," in *SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2007, pp. 434–443.

- [6] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg, "Online auctions and generalized secretary problems," *SIGecom Exch.*, vol. 7, no. 2, pp. 1–11, 2008.
- [7] M. Sakaguchi, "Dowry problems and ola policies," *Rep. Statist. Appl. Res. JUSE*, 1978.
- [8] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg, "A knapsack secretary problem with applications," in *APPROX '07/RANDOM '07: Proceedings of the 10th International Workshop on Approximation and the 11th International Workshop on Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 16–28.
- [9] P. Freeman, "The secretary problem and its extensions: A review," *International Statistical Review*, 1983.
- [10] M. A. Smith and T. Kanade, "Video skimming and characterization through the combination of image and language understanding," *Content-Based Access of Image and Video Databases, Workshop on*, vol. 0, p. 61, 1998.
- [11] S. Uchihashi, J. Foote, A. Girgensohn, and J. Boreczky, "Video manga: Generating semantically meaningful video summaries," in *Proceedings of the seventh ACM international conference on Multimedia*. ACM Press, 1999, pp. 383–392.
- [12] C. wah Ngo, Y. fei Ma, and H. jiang Zhang, "Automatic video summarization by graph modeling," in *Proceedings of the 9th IEEE International Conference on Computer Vision*, 2003, pp. 104–109.
- [13] Y. fei Ma, L. Lu, H. jiang Zhang, and M. Li, "A user attention model for video summarization," in *Proceedings of ACM Multimedia*, 2002, pp. 533–542.
- [14] Y. Gong and X. Liu, "Video summarization using singular value decomposition," in *Proc. of CVPR*, 2000, pp. 174–180.
- [15] S. M. S.X. Ju, M.J. Black and D. Kimber, "Summarization of videotaped presentations: Automatic analysis of motion and gesture," in *IEEE Transactions on Circuits and Systems for Video Technologies*, 1998.
- [16] Y. Takeuchi and M. Sugimoto, "User-adaptive home video summarization using personal photo libraries," in *Proceedings of the 6th ACM international conference on Image and video retrieval CIVR 07*, 2007.
- [17] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," in *In Proc. ICCV*, 2003, pp. 1470–1477.
- [18] F. Schaffalitzky and A. Zisserman, "Automated scene matching in movies," in *In Proceedings of the Challenge of Image and Video Retrieval, London, LNCS 2383*. Springer-Verlag, 2002, pp. 186–197.
- [19] R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations*, R. E. Miller and J. W. Thatcher, Eds. Plenum Press, 1972, pp. 85–103.
- [20] Y. Girdhar and G. Dudek, "Optimal online data sampling or how to hire the best secretaries," in *Canadian Conference on Computer and Robotic Vision(CRV)*, Kelowna, British Columbia, May 2009.
- [21] L. Itti and P. Baldi, "Bayesian surprise attracts human attention," *Vision Research*, vol. 49, no. 10, pp. 1295 – 1306, 2009, visual Attention: Psychophysics, electrophysiology and neuroimaging.
- [22] S. Kullback, *Information theory and statistics*. John Wiley and Sons, NY, 1959.
- [23] J. Sattar, G. Dudek, O. Chiu, I. Rekleitis, P. Giguère, A. Mills, N. Plamondon, C. Prahacs, Y. Girdhar, M. Nahon, and J.-P. Lobos, "Enabling autonomous capabilities in underwater robotics," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, Nice, France, September 2008.



Fig. 4. Street View Dataset 1. a) 64 images from our street view data set. b) Result of running the off-line summary algorithm, requesting eight summary images. Here the first image corresponds to the mean appearance of the scene. c) Result of running the on-line summary algorithm. We see that 4/5 images in the selected set are either same or similar to the off-line selection set.