Table 2: Human evaluation results

Written	GR	NR	RC	TF	SC
SpOpt-Δ	4.20±0.70	3.53±0.62	3.93±0.93	3.80±0.87	3.40±0.66
SpOpt-comp-Δ	3.43±0.84	3.97±0.60	4.00±0.86	3.90±0.79	3.73±0.68
Human	4.97±0.18	4.93±0.25	5.00±0.00	4.93±0.25	4.93±0.25

Table 3: Example compressive summary from DUC 2006

SYSTEM OUTPUT For many parts of the Americas and Asia, the occasional warming and cooling cycles in the tropical Pacific Ocean known as El Nino and La Nina are unwelcome visitors barging in, usually with little warning, then staying for months or sometimes a year or two, bringing all kinds of baggage in the form of distorted patterns of storms and droughts, heat and cold. Although La Nina is already affecting rainfall in Southeast Asia, it hasn't noticeably influenced North America's weather yet, Leetmaa said Instead, it may be that warm water left over from El Nino is contributing to the formation of more tropical storms than normal in the eastern Pacific. Scientists cautioned that like its warm counterpart, El Nino, a La Nina condition will influence global climate and weather until it has completely subsided. La Nina, the assertive sister of last year's hellacious El Nino, has already contributed to freakish weather around the globe and will continue to bring colder than normal temperatures to the West Coast well into spring, according to government forecasters. With no El Nino or La Nina, there is no sheep dog driving the sheep, no strong organizing force giving shape to the otherwise chaotic flow of weather across the seas and continents. La Nina and El Nino form the opposite ends of the same climatic cycle, with El Nino fueled by unusually warm water in the eastern Pacific and La Nina driven by cold. The pool of cold Pacific water that is the heart of La Nina has shrunk to one-fourth the size it was in May, even as warm water left over from El Nino remains, said Bill Patzert, a researcher at the National Aeronautics and Space Administration laboratory in suburban Pasadena.
REFERENCE SUMMARY El Nino is a disruptive weather phenomenon that usually occurs every three to four years on the average. It involves a surface warming of the eastern and central Pacific Ocean around the equator when trade winds weaken. It can disrupt climate around the world, producing extra rain in the southeastern U.S., Peru and Ecuador during the winter, while causing drought in the western Pacific as well as the slowing of trade winds and changes in sea levels. Basically, it reverses normal weather patterns, resulting in drought in usually wet locations and flooding in arid areas. It also helps trigger some man-made disasters, such as forest fires, and enhances conditions that cause viral disease among humans and livestock. La Nina is the phenomenon of rapidly cooling equatorial waters in the central Pacific. It develops every several years and works in reverse of El Nino. It can last for one year causing cooler conditions in central North America and dry warm conditions in the southern states that can be just as disruptive as El Nino. On the positive side, El Nino can be credited with saving lives that would have been lost in normal winter and hurricane seasons. It may also help cut global warming by temporarily stemming release of carbon dioxide from the Pacific Ocean. Computer module studies and satellite systems allow for a better understanding of how El Nino and La Nina form but, unfortunately, when they will develop or what they hold for the future still cannot be predicted.

5 Related Work

Our sparse optimization formulations are closely related to data reconstruction for document summarization. The data reconstruction paradigm for document summarization, originally proposed by He et al. [2012], was inspired by latent semantic analysis (LSA) that utilizes singular value decomposition (SVD) to select highly ranked sentences [Gong and Liu, 2001]. Nonnegative matrix factorization has also been introduced to group sentences into clusters [Wang *et al.*, 2008]. Recently Liu et al. [2015] propose a two-level sparse representation model. Their optimization problem is NP-hard so heuristic methods such as simulated annealing has been used to solve it approximately.

In recent years some research has made much progress beyond extractive summarization, especially in the context of compressive summarization. An earlier attempt made by Zajic et al. [2006] tried a pipeline strategy with heuristics to generate multiple candidate compressions and extract from this compressed sentences. Berg-Kirkpatrik et al. [2011] created linear models of weights learned by structural SVMs for different components and tried to jointly do sentence selection and syntax tree trimming in integer linear programs. Woodsend and Lapata [2012] designed quasi tree substitution grammars for multiple rewriting operations. All these methods involve integer linear programming solvers to generate the final compressed summary, which is time-consuming for multi-document summarization tasks.

Almeida and Martins [2013] formed the compressive summarization problem in a more efficient dual decomposition framework. Models for sentence compression and extractive summarization are trained by multi-task learning techniques. Wang et al. [2013] explored different types of compression on constituent parse trees for query-focused summarization. In these works, the best-performing systems require supervised learning for different subtasks.

Our mathematical formulations are closely related to modern sparse optimization problems. Subspace clustering techniques [Elhamifar and Vidal, 2013] try to learn proper coefficients, aiming at a self-representation. The difference between general sparse subspace clustering problems and our formulations will make key impact on the choice and design of solving algorithms. The difference comes mainly from different motivations. The former expect for sparsity in sentence selection, while the latter typically requires low-rankness in matrix representations.

6 Conclusion and Future Work

In this paper we propose a new formulation for document summarization via sparse optimization with decomposable convex objective function and derive an efficient ADMM algorithm to solve it. We also introduce a sentence dissimilarity term to encourage diversity in summaries. Then we generalize the proposed method to compressive summarization and derive a block coordinate descent procedure along with recursive dependency tree compression to generate the final sentences. Experimental study shows that our compressive summarization framework significantly improves results from the original extractive methods based on data reconstruction.

Structured sparsity has been studied for a while in machine learning community. However, its adaptation in natural language processing and text mining is still at its beginning [Martins *et al.*, 2011; Yogatama and Smith, 2014]. We would like to explore if structured sparsity can become useful for compressive summarization tasks.

Our proposed methods are fully unsupervised. We would like to extend it to supervised case for different optimization problems described in this paper. It is reasonable to expect for even better performance.

Acknowledgments

We thank all the anonymous reviewers for helpful comments and suggestions. This work was supported by National Hi-