DEWARPING CURVED DOCUMENT IMAGES WITH MATLAB

CHI ZHANG, COURSE PROJECT FOR

CPSC 635 - IMAGE ANALYSIS AND COMPUTER VISION - FALL 2017



INTRODUCTION: DIGITIZING DOCUMENTS WITH CELL PHONE CAMERA



Digitizing books



Foreign language translation



Archiving rare documents

DISTORTED IMAGE

37. DISCUSSION

However, the situation was found to be different when generalizing to domain However, the situation was larger to domain unseen in training. Here, significant overfitting was a key problem for all systems unseen in training. Here, significant overfitting was a key problem for all systems under the detectors trained on web video performed about as many unseen in training. Here, and surprisingly detectors trained on web video performed about as well as the ones and surprisingly detectors trained data. Also, detection rates could be appointed by appointed the country of the country and surprisingly detections at the ones trained on strongly annotated data. Also, detection rates could be improved in training on strongly annotated data. trained on strongly announced in this situation by supplementing conventional training sets with YouTube content overall, these results demonstrate that web video is a highly interesting data. Overall, these states of the source for concept detector training. With large-scale readily annotated data of source for concept detection systems can be trained under fered by services like YouTube, concept detection systems can be trained under less supervision, can scale up to more concepts, and thus provide better sunner for video search. Compared to the proposed web-based concept learning, a manual annotation of training sets may not really be worth the effort, as it only gives improvements on the restricted training domain. For a practical application in which a concept detector is applied to video sources unseen in training, it seems preferable to automatically bootstrap detection from web video and then performs light-weight manual refinement on the target domain, for example using relevance feedback [RL03].

OCR RECOGNIZED TEXT

trained 011

_ A1 , detection rates co ld ' gly annotated data S0 u be Improvedin

1t..t.

plementing conventiona raining se s with YouTube content.
ults demonstrate that Web Video IS a highly interesting data
tector training. With large-scale readily annotated data0f_
d by Services like YouTube, concent detection systems can be

d by Services like YouTube, concept detection systems can be trained under supervision, can scale up to more cOI1C€3Pl'{Sa and thus Provide better Support for video search. Compared to the proposed Web-based concept learning, a man. ual annotation of training sets rnayrnot really be Worth the effort, as it only gives improvements on the restricted training domain', For a practical application which a concept detector is applied to 'video sources unseen in training, it seems trai

. - u

this situation by 5 P Overall, these res

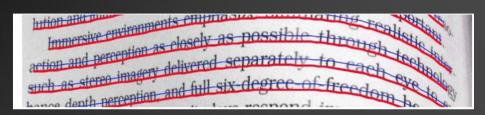
BACKGROUND: METHODS

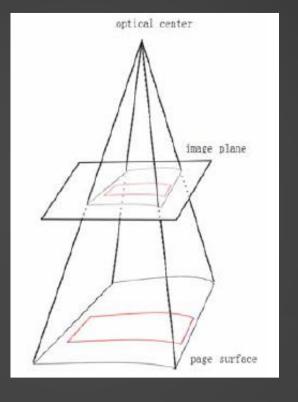
4

Warp: to bend or twist out of shape, especially from a straight or flat form Dewarp: reverse process of warp

Important emphasize simulating realistical indication and perception as closely as possible through technique such as stereo imageny delivered separately to each eye hand such as stereo imageny delivered separately to each eye hand such as stereo imageny delivered separately to each eye hand such as stereo imageny delivered separately to each eye hand such as stereo imageny delivered separately to each eye hand such as stereo imageny delivered separately to each eye hand separately experience to exper

Stroke analysis





The SB-tree of P. E. O'Neil [Acta Inf. 19 144 And Inf. 19 144

Baseline detection

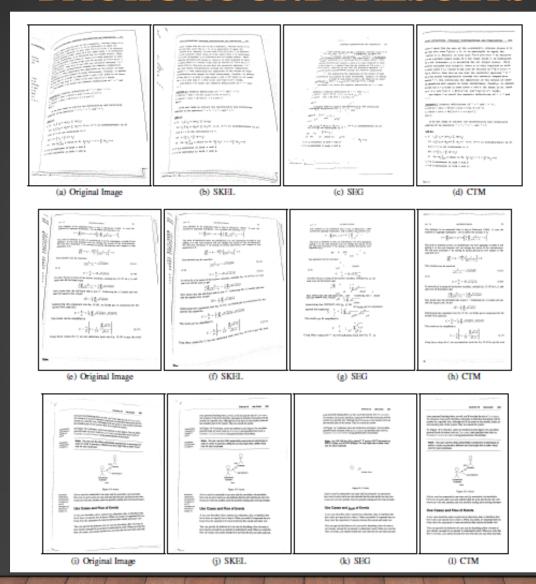
3D transform model

Brightness contrast

Y. Tian and S. G. Narasimhan, "Rectification and 3D reconstruction of curved document images," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 377–384, 2011.

B. Fu, M. Wu, R. Li, W. Li, Z. Xu, and C. Yang, "A model-based book dewarping method using text line detection," Proc. 2nd Int. Work. Camera ..., pp. 63–70, 2007.

BACKGROUND: RESULTS



Best reported results: OCR accuracy 95-100%

METHODOLOGY (1)

- Objective: dewarp document images with Matlab implementation; no external libraries
- IUPR 2011 Dataset (images of warped and scanned documents)
- http://didcontest2011.blogspot.ca/
- Image dimensions: 2592 pixels x 3456 pixels
- Assumptions
 - 1. images of individual pages from book
 - 2. focal plane almost normal to page surface
 - 3. text areas are in upright orientation (preprocessed through rotation)

METHODOLOGY (2)

Step 0. PREPROCESSING: Complement and erode (remove salt and pepper noise) Step 1. Dilate horizontally and erode vertically to turn each textline into a Refinement connected component Step 2. Connected component analysis; textline fitting with 3rd order polynomial Step 3. Text area left/right border fitting with 2nd order polynomial Refinement Step 4. Calculate intersection points between textlines and left/right borders Step 5. Calculate reference points (local min/max of textlines, borders) Step 6. Generate grid points (moving points)

Step 8. Polynomial transformation

Step 7. Generate rectified grid points (fixed points)

STEP 0. COMPLEMENT AND ERODE TO REMOVE SALT AND PEPPER NOISE

8 BEFORE AFTER

3.7 Discussion

Hosener, the situation was found to be different when generalizing to isome moral is training. Here, significant overfitting was a key problem for all spating moral is training. Here, significant overfitting was a key problem for all spating and single detection rates could be improved in a single of a country outcomed data. Also, detection rates could be improved in a single of a country supplementing conventional training sets with YouTube court; it similarly supplementing conventional training sets with YouTube court; it similarly interesting day source for concept detector training. With large-scale readily annotated data; sets by services like YouTube, concept detection systems can be trained new supervision, can scale up to more concepts, and thus provide better supple in videa search. Compared to the proposed web-based concept learning, a new manifestation of training sets may not really be worth the effort, as it only give in provenents on the restricted training domain. For a practical application which a concept concept is applied to video sources unseen in training, it is provened to the proposed when the video and then practical light-weight unsumal relinement on the target domain, for example using results is ableded Theory.

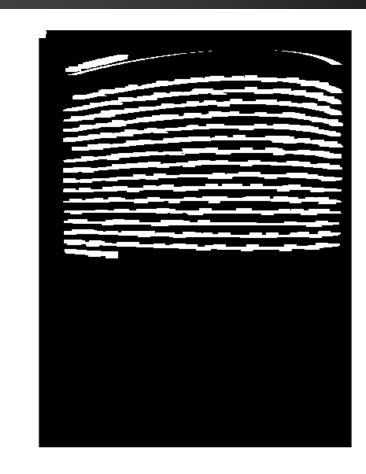
n usuud relinement on the target domain, for example using secure

Therein the advantages found in the different when generalizing to some there is no advantages from the control with video performed about as well in the analysis of the control with video performed about as well in the control with the effort, as it can be included as an including sets may not really be worth the effort, as it can be incontrol with the control with the effort, as it can be incontrol with the effort, as it can be incontrol with the control with the effort, as it can be described as a control with the effort, as it can be control described with the control with the effort, as it can be control described with the control with the

STEP 1. DILATE WITH HORIZ. ELEMENT / ERODE WITH VERT. ELEMENT

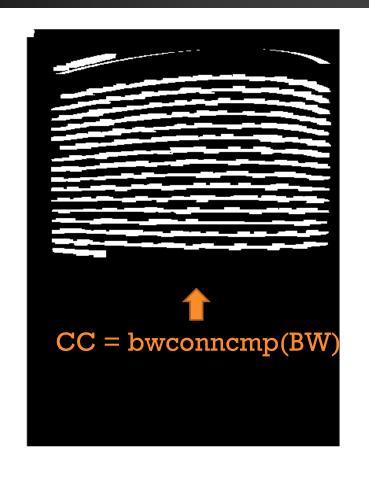
9 BEFORE AFTER

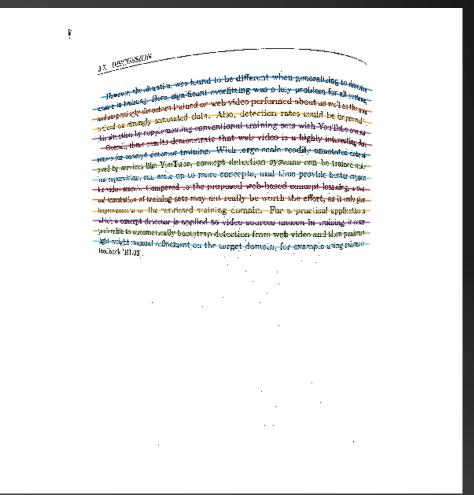
There is the above was found in the different when generalizing to some product the above was found in the above was a key problem for all ways and make the decision of the solution and the solution with the product of the colors whence on the above activated case. Also, discertion rates would be impossible to a straight continues are controlled activated case. Also, discertion rates would be impossible to the above activated case. Also, discertion rates would be impossible and the above activates activated the training sets with YorlYbe considered which was the results of the web video in a highly interesting the solution of control decision of rating. With large-scale readily aminested days solve the Yorl'unc, concept detection systems can be trained as approach, or while you would thus provide enters and a supercolor, or while the two represents with based control learning and really not would be affect, as it constitutes of the estimated the interior of worth the affect, as it makes a the estimated to vicine sources unswer in training it are above to example decision is another to sources unswer in training it are above to example decision is another to source unswer in training it are above to example decision is another than the control decision in the categor domain. For a principal infinite plant above the approach of the categor domain on the categor domain for example using search below. Blog.



STEP 2. CONNECTED COMPONENTS ANALYSIS; TEXTLINE FITTING WITH CUBIC (3RD ORDER) POLYNOMIAL

10 BEFORE AFTER





STEP 3. LEFT, RIGHT BORDER FITTING WITH QUADRATIC (2ND ORDER) POLYNOMIAL

11 BEFORE AFTER

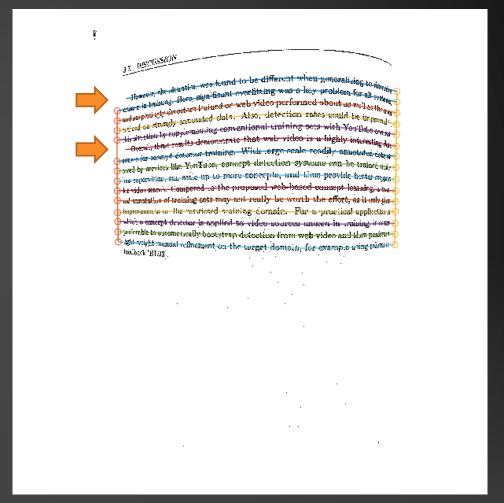
Hosein all distribution was tours to be different when generalizing to inven-House, the death wind or web rides performed shout as a line with the core is trained at a line of trained at a line of the core is trained at a line of trained at a more in the control of the control o undourprise to the property of the country of the c his straight appropriate conventional training sets with You labor on the this stration by copper and the demonstrate that web video is a highly interesting to over, the overal enter we training. With organizate readily abbotion countries and by arrices like YenTube, correspt detection systems can be trained as in appreciate, our sense up to more concepts, and time provide bester store he vilor second Compered to the proposed web located concept leaving a tw -ut contains of treesing cots may and really be worth the effort, as it only six improvements on the restricted straining domain. For a procitical applications which a omegat detector is applied to video accurees autocen in training it me-- sermble to wrome heally bootstrap detection from web video and then points. ight wight manual refinement on the target domeout, for example using more

O Beering the distribution was transfer or the different when problem for a light the contract of the contract Here's he destract we fitting was a key problem for all many trace is trained or web video performed should be made a language for the many trace. motion de la company de la com Contemporary to the mounted duty. Also, detection rates could be improved oriced on structure and the conventional training sees with You labourne this stration by supplied the demonstrate that web vicks is a highly intensiting the December to the training With organical readily absolute court Durel by services blee VanTuse, convept detection systems can be trained at has specified, our same up to more concepts, and hims provide better agenht vilor seen. Compered to the proposed web-based concept learning the at contain of treesing cots may and really be worth the effort, as it only size improvements on the restricted training domain. For a practical applications 🕩 which a concept describe is applied to video appropriate function in training it me Фэнтыв to wnome treatly bootstrap detection from web video and the points (🕁 ight-wight minus trefinement on the target donors, for example using mound

STEP 3A. REFINING LEFT BORDER (REMOVE INDENTATION)

12 BEFORE AFTER

O Hosevit almost a way tenned to be different when generalizing to distribute the different when generalizing to distribute the different when the problem for all the different when the problem for all the different when the different when generalizing to distribute the different when generalizing the different generalizing the different generalized the differen Hosein de desta mind or web rides performed shout and a start of the s more in tracers there takes are video performed about to melastic them. Contemporary Contemporary and Cuts. Also, deterritors rates could be improved. his straight attended to the well violes in the ball to the this in the chart to demonstrate that web victor is a highly interesting to Ocean the second training. With organization readily appointed course Description of the VenTube, correspt detection systems can be trained by action of the VenTube, correspt detection systems can be trained as on supervision, our sense up to more concepts, und time provide bester store he vitor section. Compered to the proposed web-based concept learning and at contains of treiting cots may and really be worth the effort, as it only size improvements on the restricted training domain. For a practical applications (🕽 🕬 i guides. W resear a correct of the field of the fi promise to enometically bootstrap detection from web yideo and the posts of 🕁 ign-while manual refinement on the turget domain, for example using more 🕙



STEP 4. INTERSECTION POINTS BETWEEN LEFT/RIGHT BORDERS AND TEXTLINES

13 BEFORE AFTER

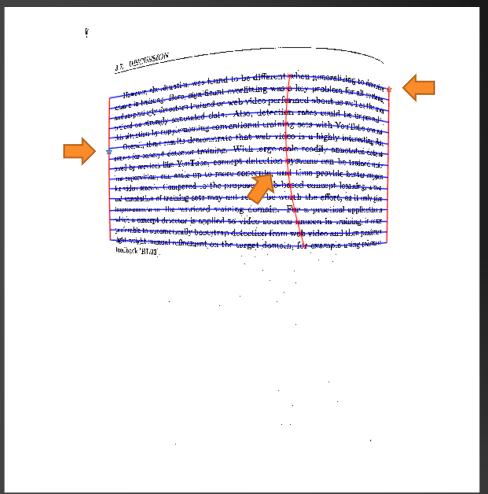
Hoseware density was tourned to be different when generalizing to imply Hoseling the details were the relative was a key problem for all water to the relative performed about to Come is traited. Here sured or web video performed about so well the medical training of training of the medical training of t and our frame of the information his direction by suppressions in a conventional training sets with You libbourne his art time by support Bearing and that web victor is a highly interesting by Ocea-the server training. With organish readily abusting the Dured by services blee VanTuran, connept detection systems can be trained in posed of white our sense up to more correcpla, and time provide bester store he vitor seems. Compered to the proposed web-based concept learning and at contains of treiting costs may not really be worth the effort, as it only size improvements on the restricted statisting domain. For a practical application (🕽 🖘 ii guute. 🙀 meenn ueeraaa oebiya ta valqaa e takeen in samue. 🖟 promise to enometically bootstrap detection from web yideo and the posts of digit wint mount refinement on the target domests, for example using mount of

Mose the situation was found to be different when generalizing to the one on its trained. Here, against on web video performed about us well so than a description rates could be a thought and represent supersted date. Also, detection rates could be improved; real or receipt of the super ing conventional training sets with YouTule cont three, these results depond to the true video is a highly interesting to thresh, three resummers, with large scale readily annotated design word by services like You'labo, concept detection systems can be trained us eren by service can seake up to more concepts, and thus provide better support for video search. Compared to the proposed web-based concept learning sucand enquestion of training sets many not really be worth the elfort, as it out me in provements on the restricted training dornain. For a practical application is which a manager concerns to applied to video courses unacon in training it was infamble to automatically bootstrap detection from web video and then paken klight weight manual relinement on the target domain, for example using surant *

STEP 5. CALCULATE REFERENCE POINTS (LOCAL MIN/MAX OF TEXTLINES, LEFTMOST POINT ON LEFT BORDER, RIGHTMOST POINT ON RIGHT BORDER)

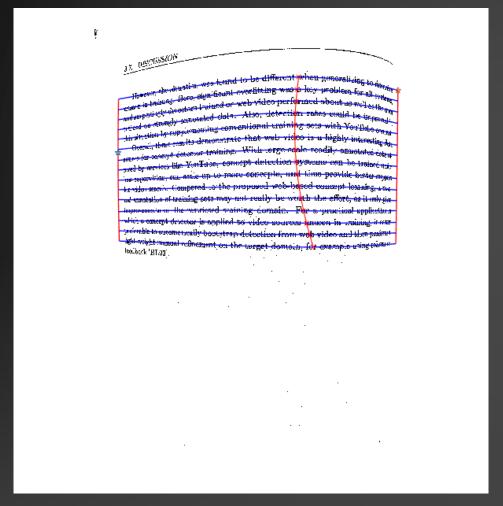
14 BEFORE AFTER

Moreover situation was found to be different when generalizing to de × Insert in the site of the significant everfitting was a key problem for all ones, and one of the site of the sit and a fair deleter trained on web video performed about us well as then and surface to complete of the Alan, detection rates could be improved ter who by supposed in conventional training sets with You Tube comp three, these results demonstrate that web video is a highly interesting to threat, these restriction training. With large scale readily announced the work for services like You Tube, concept detection systems can be trained and ere more white, can seele up to more concepts, and thus provide better sund for video states. Compared to the proposed web based concept learning our and enquestion of training sets many not really be worth the effort, as it out one in provements on the restricted training dornain. For a practical application which a manager executor is applied to video sources unacen in training it was incombie to automatically bootetrap detection from web video and then parker klight weight manual relinement on the target domain, for example noing recess



STEP 6. GENERATE MOVING POINTS (INITIAL GRID POINTS) ON TEXTLINES

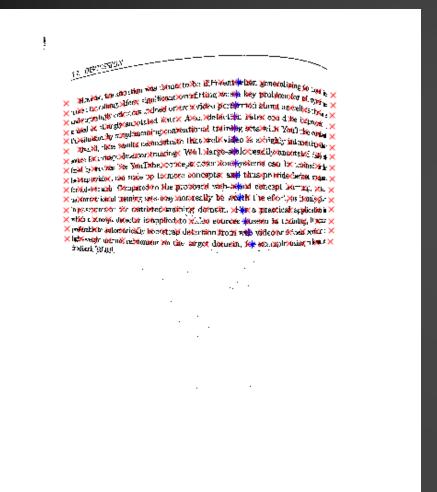
15 BEFORE AFTER



Manager in the control of the contro * Market top and then was a control or of thing we say bey produced to a to a control of a contr * The the characters and who were a vide a period at the second of the ex-X under mellige color con the state of the color of the c × c wind by their ground manner upcontraction and their white sets with Youth the ones × their winds of the contract of the co × Costuden by sugarantees released the coretx side is excitably intracting × x the M. the senter release to the coret in the coret in the coret in the core in the × New M. 1822 Section Naturality Vol. 1. derge-adolectes illuminantes in a x X south Sections (Sec Value Date) Control of Control o X forful-we and Occupated on the proposed with believed concept the man, No. X × sexument ional manning sales does manufacility by societies the effort his industrial × nposyconder on detribuctarditing dampair, xi-as a precioal sopication × X which a X resign objection is complicated to X. The endurate Museum has a committee if your X 🗙 profeszbete automárickégy bedest sajá destartánn átroda rede vádledész ádadál szású : 🗙 × latik-velajn ujetniš režinimatin žini (žije žingos) dokuešin, je≱ ses (zijalnivisias) ksp. ×

STEP 7. CALCULATE FIXED POINTS (MODIFIED GRID POINTS)

16 BEFORE AFTER



to stocking the significant even found to be difference when generalizing to the significant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for a possible fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for an applicant even fitting was to key problem for a possible fitting was to key problem for a possible fitting was to key problem for a possible fitting was to key problem. the state of the significant weeks the problem for all systems and the state of the ones in turing. Here, assumed ones video performed about sawellas them, average in a detector trained ones. Also, eletection rates could a large trained on the Also, eletection rates could a large trained on the Also, eletection rates could a large trained on the Also, eletection rates could a large trained on the Also, eletection rates could a large trained on the Also, eletection rates could a large trained on the Also, eletection rates could a large trained on the Also, eletection rates could be a large trained on the Also, eletection rates could be a large trained on the Also, eletection rates could be a large trained on the Also, eletection rates could be a large trained on the large trai and supplied by detection countries and detection and the countries of the their union by supplience sing conventional training sets with YouTube only. restruction by supplementary that well video is addigably interestingly by some is content tretestor trainings. With Marge-scale rendity annotated date; * some in college in the YouTube, concept detection by stems can be trained attent * end supportision, can scale up to more concepts: and thus provide better support * Antiderscarch. Compared to the proposed was based concept learning sur. × nalegomerations of training sets transport areally be worth the ellectras incomprises x improvements on the restricted straining dornain. XFex a practical application i X × which a wincept concentr is applied to widen sources wascen in training it was × * present to setomatically brotatrap detection from with videoxand their paters × × light-wight openual relinement on the target domain, for exampleousing record × y(x) = cubic polynomial fit to textline

RESULTS (1): DEWARPED IMAGES

17

tform = fitgeotrans(movingPoints, fixedPoints, 'polynomial', N)

Original

However, the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was found to be different when generalizing to dollaging the situation was a key problem for all situations and the situation was a situation when generalizing the situation was a situation with the situation will be situation with the si

unseen in training. Here, unseen in training. Here, unseen in training. Here, unseen in training detectors trained on web video performed about as well as the one and surprisingly detectors trained data. Also, detection rates could be expectated data. and surprisingly detections at the one trained on strongly annotated data. Also, detection rates could be improved in training conventional training sets with W trained on strongly annotated that the training sets with YouTube content this situation by supplementing conventional training sets with YouTube content is situation by supposed the content of the content Overall, these scales readily annotated data of source for concept detector training. With large-scale readily annotated data of source for detection systems can be trained under fered by services like YouTube, concept detection systems can be trained under less supervision, can scale up to more concepts, and thus provide better support for video search. Compared to the proposed web-based concept learning, a manual annotation of training sets may not really be worth the effort, as it only give improvements on the restricted training domain. For a practical application in which a concept detector is applied to video sources unseen in training, it seems preferable to automatically bootstrap detection from web video and then performs light-weight manual refinement on the target domain, for example using relevance feedback [RL03].

However, the situation was larger than the significant overfitting was a key problem for all systems unseen in training. Here, significant overfitting was a key problem for all systems.

3.7. DISCUSSION

However, the situation was found to be different when generalizing to domain unseen in training. Here, significant overfitting was a key problem for all systems and surprisingly detectors trained on web video performed about as well as the one trained on strongly annotated data. Also, detection rates could be improved in this situation by supplementing conventional training sets with YouTube content

Overall, these results demonstrate that web video is a highly interesting data source for concept detector training. With large-scale readily annotated data of fered by services like YouTube, concept detection systems can be trained under less supervision, can scale up to more concepts, and thus provide better support for video search. Compared to the proposed web-based concept learning, a manual annotation of training sets may not really be worth the effort, as it only give improvements on the restricted training domain. For a practical application in which a concept detector is applied to video sources unseen in training, it seems preferable to automatically bootstrap detection from web video and then performs light-weight manual refinement on the target domain, for example using relevance feedback [RL03].

3.7. DISCUSSION

However, the situation was found to be different when generalizing to domain. However, the statement of the However, the and surprisingly detectors trained on web video performed about as well as the one trained on strongly annotated data. Also, detection rates could be improved in this situation by supplementing conventional training sets with YouTube content

Overall, these results demonstrate that web video is a highly interesting data source for concept detector training. With large-scale readily annotated data of fered by services like YouTube, concept detection systems can be trained under less supervision, can scale up to more concepts, and thus provide better support for video search. Compared to the proposed web-based concept learning, a man ual annotation of training sets may not really be worth the effort, as it only gives improvements on the restricted training domain. For a practical application in which a concept detector is applied to video sources unseen in training, it seems preferable to automatically bootstrap detection from web video and then performs light-weight manual refinement on the target domain, for example using relevance feedback lyrace. feedback [RL03].

3.7. DISCUSSION

However, the situation was found to be different when generalizing to domain unseen in training. Here, significant overfitting was a key problem for all systems and surprisingly detectors trained on web video performed about as well as the one trained on strongly annotated data. Also, detection rates could be improved in this situation by supplementing conventional training sets with YouTube content

Overall, these results demonstrate that web video is a highly interesting date source for concept detector training. With large-scale readily annotated data of fered by services like YouTube, concept detection systems can be trained under less supervision, can scale up to more concepts, and thus provide better support for video search. Compared to the proposed web-based concept learning, a manual annotation of training sets may not really be worth the effort, as it only gives improvements on the restricted training domain. For a practical application in which a concept detector is applied to video sources unseen in training, it seems preferable to automatically bootstrap detection from web video and then performs light-weight manual refinement on the target domain, for example using relevance feedback [RL03].

N=2

N = 4

N = 3

RESULTS (2): OCR PROCESSING OF DEWARPED IMAGE

18 POLYNOMIAL TRANSFORM (N = 4)

3.7. DISCUSSION

However, the situation was found to be different when generalizing to domains unseen in training. Here, significant overfitting was a key problem for all systems, and surprisingly detectors trained on web video performed about as well as the ones trained on strongly annotated data. Also, detection rates could be improved in this situation by supplementing conventional training sets with YouTube content.

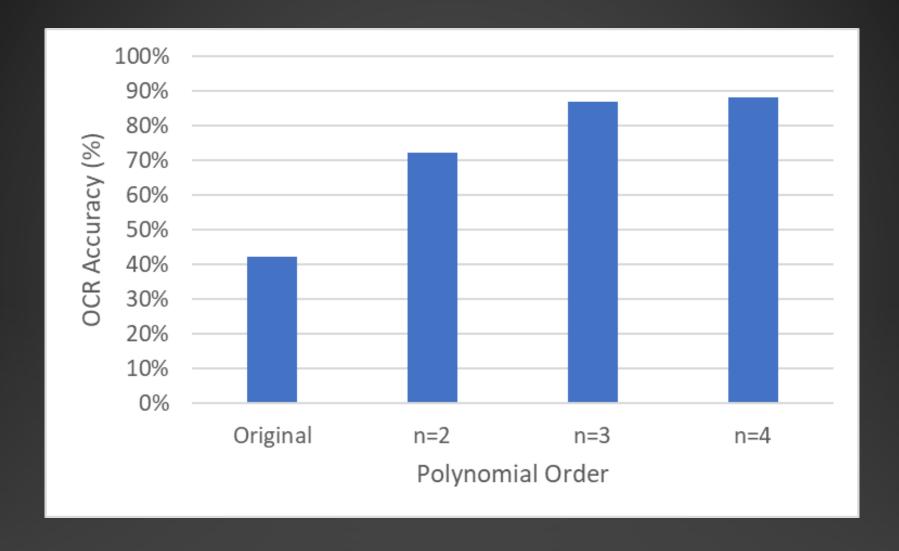
Overall, these results demonstrate that web video is a highly interesting data source for concept detector training. With large-scale readily annotated data of fered by services like YouTube, concept detection systems can be trained under less supervision, can scale up to more concepts, and thus provide better support for video search. Compared to the proposed web-based concept learning, a manual annotation of training sets may not really be worth the effort, as it only gives improvements on the restricted training domain. For a practical application in which a concept detector is applied to video sources unseen in training, it seems preferable to automatically bootstrap detection from web video and then perform a light-weight manual refinement on the target domain, for example using relevance feedback [RL03].

OCR RECOGNIZED TEXT

3.7. DISCUSSION

However, the situation was found to be different when generalizing to domains unseen in training. Here, significant overfitting was a key problem for all system, and surprisingly detectors trained on web video performed about as well as the ones trained on strongly annotated data. Also, detection rates could be improved in this situation by supplementing conventional training sets with _You'l'ube content.

Overall, these results demonstrate that web videos is a highly interesting data source for concept detector training. eWith large-scale readily annotated data offered by services like YouTube', tconceplit detection systems can be trained under less supervision, can scale up to more concepts, 5and{ thus provide better support for video search. Compared the proposed 'vve___b-l__lb_ased.ii"concept learning, a manual annotation of training ysetslfrnay:-notl.'réallyS be lvvorth-' the "effort, as it only gives improvements on the restricted —trainin.g domain;. For pracfitical applicationin which a concept detector' is'-applied eto:i{vfide¢ sources': yl11l_r1.'sLe'ely1 training, it seems preferable to automatically "bootstrapl detection pfrolm vveb videofand then perform lightweight manual refinementon,.the;*tvargetf for example using relevance feedback [RLO3I- s r _ g v =



DISCUSSION

- Run time: ~10 seconds per image
- Top OCR Accuracy: 88%, (n = 4)
- Best reported OCR Accuracy: 95-100% [REF]
- Requires manual adjustments to connected component analysis
- Dewarping method is least effective near left/right borders
- No line/paragraph spacing adjustment

CONCLUSION

- Most basic dewarping method is implemented from scratch using Matlab, binary image processing, vector calculus
- Better open source algorithms available in C, Matlab, Python, and possibly many others
- Chi Zhang, chi.zhang@ucalgary.ca