geoff_cox geoff_cox

WAYS OF MACHINE SEEING[1]

GeoffCox

You are looking at the front cover of the book Ways of Seeing written by John Berger in 1972. The text is the script of the TV series, and if you've seen the programmes, you can almost hear the distinctive pedagogic tone of Berger's voice as you read his words: "The relation between what we see and what we know is never settled." [2]

The image by Magritte on the cover further emphasises the point about the deep ambiguity of images and the always-present difficulty of legibility between words and seeing.[3] In addition to the explicit reference to the "artwork" essay by Walter Benjamin,[4] the TV programme employed Brechtian techniques, such as revealing the technical apparatus of the studio; to encourage viewers not to simply watch (or read) in an easy way but rather to be forced into an analysis of elements of "separation" that would lead to a "return from alienation".[5] Berger further reminded the viewer of the specifics of the technical reproduction in use and its ideological force in a similar manner: "But remember that I am controlling and using for my own purposes the means of reproduction needed for these programmes [...] with this programme as with all programmes, you receive images and meanings which are arranged. I hope you will consider what I arrange but please remain skeptical of it."

That you are not really looking at the book as such but a scanned image of a book — viewable by means of an embedded link to a server where the image is stored — testifies to the ways in which what, and how, we see and know is further unsettled through complex assemblages of elements. The increasing use of relational machines such as search engines is a good example of the ways in which knowledge is filtered at the expense of the more specific detail on how it was produced. Knowledge is now produced in relation to planetary computational infrastructures in which other agents such as algorithms generalise massive amounts of (big) data.[6]

Clearly algorithms do not act alone or with magical (totalising) power but exist as part of larger infrastructures and ideologies. Some well-publicised recent cases have come to public attention that exemplify a contemporary politics (and crisis) of representation in this way, such as the Google search results for "three black teenagers" and "three white teenagers" (mug shots and happy teens at play, respectively).[7] The problem is one of learning in its widest sense, and "machine learning" techniques are employed on data to produce forms of knowledge that are inextricably bound to hegemonic systems of power and prejudice.

There is a sense in which the world begins to be reproduced through computational models and algorithmic logic, changing what and how we see, think and even behave. Subjects are produced in relation to what algorithms understand about our intentions, gestures, behaviours, opinions, or desires, through aggregating massive amounts of data (data mining) and machine learning (the predictive practices of data mining).[8] That machines learn is accounted for through a combination of calculative practices that help to approximate what will likely happen through the use of different algorithms and models. The difficulty lies in to what extent these generalisations are accurate, or to what degree the predictive model is valid, or "able to generalise" sufficiently well. Hence the "learners" (machine learning algorithms), although working at the level of generalisation, are also highly contextual and specific to the fields in which they operate in a coming together of what Adrian Mackenzie calls a "play of truth and falsehood".[9]

Thus what constitutes knowledge can be seen to be controlled and arranged in new ways that invoke Berger's earlier call for skepticism. Antoinette Rouvroy is similarly concerned that algorithms begin to define what counts for knowledge as a further case of subjectivation, as we are unable to substantively intervene in these processes of how knowledge is produced. [10] Her claim is that knowledge is delivered "without truth" through the increasing use of machines that filter it through the use of search engines that have no interest in content as such or detail on how knowledge is generated. Instead they privilege real-time relational infrastructures that subsume the knowledge of workers and machines into generalised assemblages as techniques of "algorithmic governmentality". [11]

In this sense, the knowledge produced is bound together with systems of power that are more and more visual and hence ambiguous in character. And clearly computers further complicate the field of visuality, and ways of seeing, especially in relation to the interplay of knowledge and power. Aside from the totalizing aspects (that I have outlined thus far), there are also significant "points of slippage or instability" of epistemic authority, [12] or what Berger would no doubt identify as the further unsettling of the relations between seeing and knowing. So, if algorithms can be understood as seeing, in what sense, and under what conditions? Algorithms are ideological only inasmuch as they are part of larger infrastructures and assemblages.

But to ask whether machines can see or not is the wrong question to ask, rather we should discuss how machines have changed the nature of seeing and hence our knowledge of the world.[13] In this we should not try to oppose machine and human seeing but take them to be more thoroughly entangled — a more "posthuman" or "new materialist" position that challenges the ontoepistemological character of seeing — and produces new kinds of knowledge-power that both challenges as well as extends the anthropomorphism of vision and its attachment to dominant forms of rationality. Clearly there are other (nonhuman) perspectives that also illuminate our understanding of the world. This pedagogic (and political) impulse is perfectly in keeping with Ways of Seeing and its project of visual literacy.[14] What is required is an expansion of this ethic to algorithmic literacy to examine how machine vision unsettles the relations between what we see and what we know in new ways.

```
// Recognize text on an image using Tesseract API and print it to the screen
                           // Usage: ./tess image.png
                           #include
                           #include
                           #include
                           #include
                                                                                                              Notes
                           #include
                           #include "opencv2/core/core.hpp"
                           #include "opencv2/features2d/features2d.hpp"
                           #include "opencv2/highgui/highgui.hpp"
                                                                                                             [1] The title of this paper is taken from a workshop organized by the
                           #include "opencv2/imgproc/imgproc.hpp"
                                                                                                              Cambridge Digital Humanities Network, convened by Anne Alexander,
                           #include
                                                                                                             Alan Blackwell. Geoff Cox and Leo Impett, and held at Darwin College,
                                                                                                             University of Cambridge, 11 July 2016,
                           using namespace cv;
                                                                                                              http://www.digitalhumanities.cam.ac.uk/Methods/waysofmachineseeing
                           using namesnace std
                                                                                                             [2] Ways of Seeing, Episode I (1972), https://www.youtube.com/watch?
                                                                                                              v=opDE4VX9Kk. The 1972 BBC four-part television series of 30-minute
                                                       ¹gv)
           a Pelican Original
                                                                                                              films was created by writer John Berger and producer Mike Dibb.
                                                                                                              Berger's scripts were adapted into a book of the same name, published
                                                                                                              by Penguin also in 1972. The book consists of seven numbered essays:
                                                                                                              four using words and images; and three essays using only images. See
                                                        specify the input image!" << std::endl;</pre>
WAYS OF SEEING
                                                                                                             https://en.wikipedia.org/wiki/WaysofSeeing.
                                                                                                              [3] René Magritte, _The Key of Dreams (1930),
        Based on the BBC television series with
                                                                                                              https://courses.washington.edu/hypertxt/cgi-
           JOHN BERGER
                                                                                                             bin/book/wordsinimages/keydreams.jpg. Aside from the work of
                                                       SHEY_SIMPLEX;
                                                                                                              Magritte, Joseph Kosuth's One and Three Chairs (1965) comes to mind,
                                                       *api = new tesseract::TessBaseAPI();
                                                                                                             that makes a similar point in presenting a chair, a photograph of the
                                                       -ocr with English, without specifying tessdata pat
                                                                                                              chair, and an enlarged dictionary definition of the word "chair",
            Seeing comes before words. The child looks
                                                                                                              https://en.wikipedia.org/wiki/OneandThreeChairs.
          zes before it can speak
                                                       ild not initialize tesseract.\n");
                                                                                                              [4] The first section of the programme/book is acknowledged to be
           But there is also another sense in which seeing
 comes before words. It is seeing which establishes our place
                                                                                                              largely based on Benjamin's essay "The Work of Art in the Age of
                                                                                                              Mechanical Reproduction" (1936),
 but words can never undo the fact that we are surrounded by
 it. The relation between what we see and what we know is
                                                                                                              https://www.marxists.org/reference/subject/philosophy/works/ge/benja
                                                        n leptonica library
                                                                                                              min.htm.
                                                                                                              [5] The idea is that "separation" produces a disunity that is disturbing
                                                        /_LOAD_IMAGE_COLOR);
                                                                                                              to the viewer/reader — Brecht's "alienation-effect" (Verfremdungeffekt)
                                                                                                              — and that this leads to a potential "return from alienation". See
                                                                                                              https://en.wikipedia.org/wiki/Distancingeffect.
                                                       ls, im.type(), cv::Scalar(0));
                                                                                                              [6] To give a sense of scale and its consequences, Facebook has
                                                                                                              developed the face-recognition software DeepFace. With over 1.5 billion
                                                                                                              users that have uploaded more than 250 billion photographs, it is
                                                                                                              allegedly capable of identifying any person depicted in a given image
                                                       DAD_IMAGE_GRAYSCALE);
                                                                                                              with 97% accuracy. See
                                                        3);
                                                                                                              https://research.facebook.com/publications/deepface-closing-the-gap-
                                                        , CV_GRAY2BGR );
                                                                                                              to-human-level-performance-in-face-verification/.
                                                                                                              [7] Antoine Allen "The 'three black teenagers' search shows it is society,
                                                        1, CV_PI/180, 80, 80, 10 );
                                                                                                              not Google, that is racist", The Guardian (10 June 2016),
                                                                                                             https://www.theguardian.com/commentisfree/2016/jun/10/three-black-
 a painting called The Key of Dreams
                                                        ines.size(); i++ )
                                                                                                             teenagers-google-racist-tweet. [8] Adrian Mackenzie, "The Production
            The way we see things is affected by what we
                                                                                                            ு of Prediction: What Does Machine Learning Want?," European Journal
                                                                             of Cultural Studies, 18, 4-5 (2015): 431.
                                                                                                              [9] Mackenzie, "The Production of Prediction", 441.
                                                                                                              [10] See, for instance, Antoinette Rouvroy's "Technology, Virtuality and
                                                          Point(lines[i][0], lines[i][1]),
                                                            ][2], lines[i][3]), Scalar(0,0,255), 3, 8 );
                                                                                                             Utopia: Governmentality in an Age of Autonomic Computing", in The
                                                                                                              Philosophy of Law Meets the Philosophy of Technology: Computing and
                                                                                                              Transformations of Human Agency, eds. Mireille Hildebrandt and
                               if(numlines>40){
                                                                                                              Antoinette Rouvroy (London: Routledge, 2011), 136–157.
                                                             im.rows,im.cols,im.type());
                                                                                                             [11] This line of argument is also close to what Tiziana Terranova has
                             Resid on the BEO service series wil
                                                                                                              called an "infrastructure of autonomization", making reference to
                                   api->Recognize(0);
                                                                                                              Marx's views on automation, particularly in his "Fragment on
                                   //outText = api->GetUTF8Text();
                                                                                                              Machines", as a description of how machines subsume the knowledge
                                   tesseract::ResultIterator* ri = api->GetIterator();
                                                                                                              and skill of workers into wider assemblages. Tiziana Terranova, "Red
                                                           rlevel_level = tesseract::RIL_SYMBOL;
                                                                                                              Stack Attack! Algorithms, capital and the automation of the common",
                                                                                                              Effimera (2014), accessed August 24, 2016, http://effimera.org/red-
                                                                                                              stack-attack-algorithms-capital-and-the-automation-of-the-common-
                                                                                                              di-tiziana-terranova/.
                                                                    GetUTF8Text(level);
                                                                                                              [12] Mackenzie, "The Production of Prediction", 441.
                                                                     nce(level);
                                                                                                              [13] I take this assertion from Benjamin once more, who considered the
                                                                                                             question of whether film or photography to be art secondary to the
                                                 BoundinaBox(level. &x1. &v1. &x2. &v2):
                                                      'word: '%s'; \tconf: %.2f; BoundingBox: %d,%d,%d,%d;\nquestion of how art itself has been radically transformed:
                                                                                                              "Earlier much futile thought had been devoted to the question of
                                                    word, conf, x1, y1, x2, y2);
                                                                                                              whether photography is an art. The primary question — whether the
                                                                                                             very invention of photography had not transformed the nature of art —
                                            -i->WordFontAttributes(&ignore,&ignore,&ignore,&ignore,&ignore,&ignore,&ignore,&ignone). The film theoreticians asked the same ill-
                                                          printf("point size: %d font id: %d", psize, fid); considered question with regard to film."
                                                                                                              https://www.marxists.org/reference/subject/philosophy/works/ge/benja
                                                             = QString::fromUtf8(word);
                                                                           [14] Berger was associated with The Writers and Readers Publishing \tlength: %d\n",word,utfw.length@ooperative, aiming to "advance the needs of cultural literacy, rather
                                                          size()>4){
                                                           intf("word: '%s'
                                                           Rect=CvRect();
                                                                                                              than cater to an 'advanced' [academic] but limited readership" (From
                                                                                                              the Firm's declaration of intent). In this sense it draws upon the Marxist
                                                                                                              cultural materialism of Raymond Williams and Richard Hoggart's The
                                                                  dth=(x2-x1);
                                                                                                              Uses of Literacy (1966).
                                                                     t=(y2-y1);
                                                        letterim=im(letterRect).clone();
                                                          im.copyTo(wim(Rect(letterRect.x, letterRect.y, letterRect.width, letterRect.height)));
                           //
                                                       cv::Rect roi( cv::Point( letterRect.x, letterRect.y), cv::Size( letterRect.width, letterRect.height));
                           //
                                                          printf("x %d y %d width %d height %d \n",letterRect.x, letterRect.y,letterRect.width, letterRect.height);
                           //
                                                         printf("cols %d rows %d\n",wim.cols,wim.rows);
                                                       cv::Mat destinationROI = wim( roi );
                                                       letterim.copyTo( destinationROI );
                           //
                                                          QString qs = QString("words/word_%1_").arg(word);
                                                          qs.append(QString("%1.jpg").arg(j));
                           //
                           //
                                                          imwrite(qs.toStdString(),letterim);
                                                    catch( cv::Exception& e )
                                                       const char* err_msg = e.what();
                           //
                                                          std::cout << "exception caught: " << err_msg << std::endl;</pre>
                                                    rectangle(wim,Point(x1, y1),Point(x2,y2),Scalar(255,0,0),1,8);
                                                    putText(wim,word,Point(x1,y1),fontface,0.5,CV_RGB(255,0,0), 1, 8);
                                                    confidence_level+=conf;
```

j++;